1-1-2013

Perceptions Surrounding the Implementation of Colorado Senate Bill 10-191's New Teacher Evaluations

Sarah Melvoin Bridich
University of Denver

PERCEPTIONS SURROUNDING THE IMPLEMENTATION OF COLORADO SENATE BILL 10-191'S NEW TEACHER EVALUATIONS

_____

A Dissertation

Presented to

the Morgridge College of Education

University of Denver

_____

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

_____

by

Sarah Melvoin Bridich

June 2013

Advisor: Dr. Kent Seidel

Author: Sarah Melvoin Bridich
Title: PERCEPTIONS SURROUNDING THE IMPLEMENTATION OF COLORADO SENATE BILL 10-191'S NEW TEACHER EVALUATIONS
Advisor: Dr. Kent Seidel
Degree Date: June 2013

Abstract

In 2013, many public education reform efforts in the United States of America center on testing and accountability. Recent data revealed that teachers have the single greatest in-school impact on student learning; however, the methods to assess teacher effectiveness are widely criticized for not holding teachers accountable and, consequently, are experiencing significant legislative attention. In 2010, Colorado passed Senate Bill 10-191: The Great Teachers and Leaders Act to improve student learning by revising teacher and principal evaluations, including linking them to student learning data, and eradicating tenure.

Teachers, administrators, and policymakers hold critical roles in the implementation of this bill, yet little is known about how members of each group perceive their respective roles in the implementation. This explanatory sequential mixed methods study was designed to gather perception data from these three groups, through surveys and interviews. Data revealed that teachers and administrators do not have similar perceptions to many matters related to teacher evaluations, education reform, and the implementation of Senate Bill 10-191 (SB 191). The data also revealed that teachers and administrators expected they would agree on these matters. These collective findings led to multiple recommendations, such as the need for increased dialogue between teachers and administrators about their own perceptions of education reforms.

Key words: teacher evaluations, education reform, explanatory sequential design

Acknowledgements

A heartfelt thank you to the dedicated members of my dissertation committee: Dr. Kent Seidel, Dr. Linda Brookhart, Dr. Lyndsay Agans, and the Honorable Barbara O'Brien. Your time, insights, questions, and support were instrumental to this process and greatly appreciated.

Thank you to the many family and friends who supported this educational journey in numerous ways, though not all can be mentioned specifically. Particular gratitude to Dick Light, Sarah Lewis, and Floyd Cobb for your insights. Sameera Fazili, Evelyn Chung, Ashika Singh and Arpa Garay: thanks for the daily laughs and collective motivation. Extra appreciation for Emmi Adler, Phyllis Adler, Rachel Adler, Nancy Feldman, and Susi Landers; you helped make this degree possible through your love of the munchkins. I am so grateful to have wonderful in-laws (Rick, Mardi, Nick, Mary, & Liza), and especially appreciate Mardi's willingness to fly out multiple times over the years. Extra thanks to my sister, Becca (and her Ryan), for her love and encouragement. Mom and Dad, special thanks for your extraordinary support in every area of this endeavor, too many to name, and for your many, many good questions along the way.

Thank you, Ben and Maya, for your patience with this process. I know the "disteration" took me away from spending time with both of you, and you were great sports. Jeff, thanks for your incredible love and support, and for understanding me and my need to complete this (yet another) program, despite the demands it put on our family.

Finally, I dedicate this to my grandparents, Hugo and Lois Melvoin, who have shared their passion for learning and joy for life with our whole family. Papa, you are missed.

Table of Contents

List of Tables

Chapter 1: Introduction

Background

The history of modern education reforms in the United States waxes high on failed efforts (Sarason, 1990; Tyack & Cuban, 1995; Cuban, 1998; Tyack & Tobin, 1993; Hargreaves & Shirley, 2008). The lack of teacher and administrator input in the design of reform efforts, at both the legislative and implementation levels, is one of the primary reasons these reforms are unsuccessful (Sarason, 1990; Tyack & Cuban, 1995; Hargreaves & Shirley, 2008; Berry, 2010). As Hargreaves and Shirley note (2008), "[e]ducation leaders and teachers are the ultimate arbiters of change. The classroom door is the portal to reform or the raised drawbridge that holds it at bay" (pp. 59-60). This means that, regardless of the intentions behind a reform at the legislative level, reforms that fail to incorporate teachers' and administrators' beliefs and perspectives are generally doomed at the implementation level. Education historian Larry Cuban stated this another way: "Schools change reforms as much as reforms change schools" (1998, p. 455). To understand how reforms will fare once they are practiced inside actual schools and working with actual students, it is imperative to consult with teachers and administrators. To understand the meta-picture of education reform movements in action, it is also critical to circle back to the policymakers who passed the reform to explore

whether the stakeholders have similar understandings of successful implementation (Cuban, 1998).

Purpose of the Study

The purpose of this mixed methods study was to explore how teachers, administrators, and policymakers viewed their role in the implementation of an education reform. The education reform being studied was Colorado Senate Bill 10-191: The Great Teachers and Leaders Act (SB 191),[1] which aims to improve student learning by overhauling the teacher and principal evaluation system and eradicating teacher tenure. Within SB 191, this study focused specifically on the new teacher evaluation system. The subjects in this study were teachers and administrators in the Rockies School District (pseudonym) (RSD), a single, large school district in the greater Denver, Colorado area, and policymakers who were members of Colorado's 2013 General Assembly.

Problem Statement

Given the national focus on SB 191 (Hatlestad, 2012; New Millennium Teachers, 2012; Rich, 2012; Brill, 2011) as one of the first bills to significantly alter the modern teacher and principal evaluation processes, it was imperative to explore how key stakeholders—teachers, administrators, and policymakers—each viewed their role in the implementation of SB 191's new teacher evaluation system. The bill's attempt to achieve greater educational equity for students statewide by systematically shifting the job expectations for educators, as evidenced by the overhauled teacher (and principal) evaluation systems, places it among many national efforts to use education policy as a

[1] Three years into the passage of Colorado Senate Bill 10-191, this legislation is also now called Ensuring Quality Instruction Through Educator Effectiveness (EQuITTE) (New Millennium, 2012). For sake of simplicity, this bill will be referred to only as SB 191.

lever for change. Historically, these efforts have failed; hence the need for new reforms to seek these ends yet again. For this bill to achieve success, it could be instructive to gain insights into the perceptions from these three key stakeholder groups as well as subgroups within them.

Significance of the Problem

Although no recent report has been as incendiary as the 1983 *Nation at Risk* report (National Commission on Excellence in Education), current education reform talk is dire, mired in discussions about failures and problems (Senge et al., 2000; Ravitch, 2010; Barnett, 2011; Labaree, 2010; Brill, 2011; Kane & Staiger, 2012). Students from the United States are failing to compete internationally (Friedman, 2011; Dillon, 2010). The International Association for the Evaluation of Educational Achievement's PIRLS and T*rends in International Mathematics and Science Study* (TIMSS) 2011 exams shows that some American students are improving (Berliner, 2005a). However, Secretary of Education Duncan pronounces older students' results on such assessments as "unacceptable" (Resmovits, 2012).

The current failures of students in the U.S. to test at competitive levels when compared to students from other countries has contributed significantly to the accountability and testing movement in public education. Furthered by the No Child Left Behind Act of 2001 (NCLB), which sought to increase academic achievement for low-achieving students in high-poverty schools, public schools across the nation have been increasing the number of tests they administer to students (Forte, 2010). Most states in the 1980s required the administration of only one set of achievement tests for high

poverty areas and another (or no) test for all other communities (Forte, 2010). A result of NCLB is that all states require annual standardized assessments in reading and mathematics in grades three through eight.

Recent research supports a direct link between teacher quality and student learning (Wright et al., 1997; Rockoff, 2004; Sass, 2008; Johnson, 2009; Weisberg et al., 2009; Kane and Staiger, 2012). At the same time, it is argued that the current teacher evaluation systems are not working (Weisberg et al., 2009; Meyer, 2010; Kane & Staiger, 2012). Under current teacher evaluations, most teachers are rated as exceptional (Weisberg, et al., 2009; Kane & Staiger, 2012). Yet students perform far below the "exceptional" level on state, national and international assessments (Friedman, 2011; Dillon, 2010). Although numerous factors contribute to students' performances on said assessments, there appears to be a disconnect in the dichotomy between teacher ratings and student performance that is problematic. Precisely which elements of teaching lead to improved student learning and how to measure those elements remains unclear (Johnson, 2009; Chetty, Friedman, & Rockoff, 2012), though this is presently being explored (Kane and Staiger, 2012).

In an effort to improve teacher effectiveness, many researchers (Wright et al, 1997; McCaffrey et al., 2004; Martineau, 2006; Mariano, McCaffrey, Lockwood, 2010) have developed new methods, generally known as growth models, for evaluating teachers and schools. The most well-known of these methods is the value-added model. The overarching concept is that teachers should be evaluated based on the value they add to their students' learning. Value-added models consist of complicated algorithms designed

to determine the value a particular teacher adds to a particular student or set of students over the course of a school year (Alicias, 2005; Amrein-Beardsley, 2008; McCaffrey et al., 2004; Martineau, 2006; Mariano, McCaffrey, Lockwood, 2010). In Colorado, the growth model used allows student, school, and district learning growth to be compared to learning data from other students, schools, and districts (Bonk, 2010). The objective of Colorado's growth model is to determine relative growth. For instance, a growth model can ascertain whether a student who is officially "below proficient" on his state's 3rd grade reading test improved by 1%, 10%, or 90% compared to other students in this similar bracket; this data creates an opportunity to see if that particular student is "learning" or not during a single school year, separate from the student's official classification on statewide assessments.

The 2009 United States Department of Education's "Race to the Top" fund was designed to inspire many states to rethink their education platforms, make changes as necessary, and then apply to receive some of the $4.35 billion dollars allocated to this grant. In President Obama's words, the guiding theory was that it was "time to stop talking about education reform and time to start doing it" (U. S. Department of Education Executive Summary, 2009, p. 1). The Race to the Top (RTTT) application had numerous requirements, including requirement D: Great Teachers and Leaders. Requirement D(2) called for measuring student growth and connecting this student growth to teachers and principals through annual evaluations (U. S. Department of Education Executive Summary, 2009).

One of the states competing for RTTT funding was Colorado. Although Colorado never received RTTT grant money, the state still made significant changes to its education policy as it was applying for RTTT, largely through the passage of SB 191. The link between SB 191 and RTTT is apparent just from the full name of Colorado's bill: Colorado Senate Bill 10-191: The Great Teachers and Leaders Act. The connections do not end there, as SB 191 completely modified the teacher and principal evaluation processes in ways delineated by the RTTT Fund application requirements, such as parts D(2)(ii) and D(2)(iii) which mandated "rigorous, fair, and transparent evaluation systems for teachers and principals" that used multiple rating categories and were conducted annually, respectively (U. S. Department of Education Executive Summary, 2009, p. 9).

The passage of SB 191 was highly controversial (Haley, 2010; Meyer, 2010c; Barnum, 2012; Meyers, Barnum & Fender, 2010; Pena & Zeller, 2010), illuminating a growing national schism between Democratic politicians and one of their most loyal supporters of the past few decades: teachers unions (Meyer, 2010b; Brill, 2011; Ooms, 2011). The lead author of the bill, then-freshman Democratic Senator Johnston, gathered a large and diverse coalition of supporters who collectively both supported the successful passage of the bill and defeated the Colorado Education Association (CEA)'s numerous attempts to kill the bill. The bill also benefitted from the unanimous support of Colorado's Legislative Assembly's Republicans. Despite the CEA's staunch opposition to the bill, its leadership quickly changed positions after SB 191 became law. The CEA had two positions on the 15-person State Council for Educator Effectiveness (SCEE) and

worked with coalition members to help design the specific measures of the new evaluations.

During the 20 months between the SCEE's proposal of recommendations to Colorado's State Board of Education (SBE) and the General Assembly's final vote on SBE's recommendations, the two issues that generated the most extensive debates were the number of potential final scores receivable on the teacher and principal evaluation matrices and the issue of who was designing the evaluations (Johnston & Barker, 2011). In the final passage of the bill, the evaluation for teachers and principals was to have three scoring options: Highly Effective, Effective, and Ineffective (Senate Bill 10-191, 2010). The SCEE proposed four options, but there was significant debate around the inclusion of a fifth middle option while the SCEE's recommendations were with the SBE. Ultimately, the four-option matrix was selected, with the teacher evaluation rubric rating teachers in one of these four categories: Exemplary, Accomplished, Proficient, and Partially Proficient (Colorado Department of Education, 2012). The other issue that generated debate was that of districts being expected to develop their own matrices while being offered the opportunity to *opt in* to a statewide evaluation matrix or being required to file to *opt out* of the statewide matrixes, should its leaders believe a different system to be better for the district. In essence, this second matter was one of local control and one that districts fought hard to maintain—and they won (Robles, 2011).

With these and other early implementation matters resolved, the implementation of SB 191 progressed to the pilot phase. In the 2012-2013 school year, elements of SB

191 were piloted in 27 school districts across Colorado with even more districts eager to participate (Engdahl, 2011).

In spite of the sustained and significant effort to transform SB 191 into a practice that will improve student learning around the state, the long-term implementation of SB 191 is still undefined. Policymakers were responsible for outlining the change goals through the law, but the individuals responsible for enacting it are educators on the ground level. It is unclear if teachers, administrators, and policymakers hold similar ideas regarding successful implementation. Similarly, it is currently unknown how these three key stakeholders perceive their role in the implementation of SB 191.

Enacting lasting changes in school is a complicated endeavor; change theory is an entire school of thought. Given the poor historical success of the United States's recent education reform efforts, it is critical to learn how members of each stakeholder group view their role in the implementation process. Attaining such an understanding may shed important light on the change processes involved as well as the hopes and needs of the various players. It is essential to develop an understanding of how members of each stakeholder group feels and to explore potential implementation modifications based on these results to avoid having SB 191 end in an education reform graveyard.

Research Questions

The research questions for this study were:

1) How do Colorado's teachers, administrators, and policymakers perceive their respective roles in the implementation of SB 191's new teacher evaluations?

2) Do Colorado's teachers', administrators', and policymakers' responses vary by demographic variables?

3) How do the responses of Colorado's teachers, administrators, and policymakers compare?

Definitions of Terms

For the purposes of this paper, key words and concepts will be defined as follows:

- *Administrator*: school-based position of authority, such as a principal or assistant principal. District-level administrators will be referred to explicitly as district-level administrators.
- *Colorado Growth-Model (CGM)*: a statistical model that focuses on percentile growth rather than raw results.
- *Non-probationary status*: See Tenure.
- *Organization for Economic Co-operation and Development (OECD)*: Developed in 1947, this international organization is based in Paris, France and seeks to advance policies that will improve economic and social conditions of individuals around the world.
- *Post-positivism*: a worldview that evolved from the positivist paradigm; it focuses on a single reality.
- *Programme for International Student Assessment* (PISA): the OECD's international test developed in 2000 that tests 15 year old students in three subject areas (reading literacy, mathematics literacy, and science literacy).
- *Race to the Top* (RTTT): a funded competition created by the U.S. Department of Education to distribute $4.35 billion to states that proposed models for changing their methods of conducting public education in a way that matched the competition's guidelines.

- *State Board of Education* (Colorado's)(SBE): a seven person board, with one member from each of Colorado's seven congressional districts; participation is unpaid and members have six year terms.
- *State Council for Educator Effectiveness* (SCEE): a 15 person council created by former Governor Bill Ritter to help design the early transformation of SB 191 from a bill to an active policy.
- *Teacher*: a licensed individual who works with students in a school setting; for this study, the term includes classroom instructors as well as "other licensed professionals," as determined by the CDE to include: school audiologists, nurses, occupational therapists, physical therapists, psychologists, social workers, speech-language pathologists, counselors, and teachers on special assignment (Educator Effectiveness FAQs, 2012).
- *Teacher evaluations*: method of evaluating a teacher's effectiveness; can be both formative, supporting growth, and summative, assessing job quality.
- *Tenure*: position of job security once held by Colorado's teachers following three years of satisfactory teaching. This was also known in Colorado as non-probationary status.

Summary

Education reforms abounded in the United States in 2013. Given the bleak history of modern education reforms, it may be useful for those implementing SB 191 to keep the perspectives of three key stakeholder groups—teachers, administrators, and policymakers—in the forefront of implementation design in order to increase the chances

of SB 191 meeting its official objective: improving learning for all students across the state (Senate Bill 10-191, 2010; Johnston, 2010).

Chapter Two: Literature Review

The purpose of this study was to understand how teachers, administrators, and policymakers perceived their respective roles in the implementation of SB 191's new teacher evaluations. This chapter reviews and discusses teacher effectiveness and evaluation reforms within the larger testing and accountability reform movements, the historic intractability of certain elements of public education, key findings from the organizational change and sustainable leadership traditions as they relate to systemic changes in schools and, finally, critical social theories as the framework for this study.

Teacher Effectiveness and Evaluations

Although numerous out-of-school elements impact student learning and testing (Paul, 2012), it is now widely presumed that, within schools, teachers are the single most important factor (Chetty, Friedman, & Rockoff, 2012; Kristof, 2012; Johnson, 2012; Weisberg et al., 2009). Teachers can, and do, impact students' long-term life choices, from when to have a baby to how much money one earns over a lifetime (Chetty, Friedman, & Rockoff, 2012). Despite data illustrating that teachers are the most important in-school influence on student learning, precisely which characteristics make someone a strong teacher, or in today's verbiage an "effective" teacher, remains unclear (Chetty, Friedman, & Rockoff, 2012). Currently, many states have a de facto answer in their teacher evaluations, for, as the old adage goes, what is measured is what counts.

Teacher evaluations in the USA hold a dual purpose: formative and summative assessment. The formative part consists of evaluating teachers for the purposes of improving practice: pedagogy, classroom management, approaches to content mastery, etc. The summative part is job review and determination. In other words, the formative element is designed to improve a teacher's craft, while the summative piece focuses on whether or not a teacher is qualified to have a job.

From an historical perspective, teacher evaluations are an accepted part of the teaching profession. American teachers have been evaluated, in one form or another, since the 1600s (Robles, 2007). In the late 1800s, teacher evaluation took on a particularly appearance-heavy focus, with administrators concerned about how teachers presented themselves and then how they disseminated information (Robles, 2007). Within the first two decades of the 1900s, teacher evaluations became more objective and a ratings system was introduced (Robles, 2007). Although the ratings system of the early 1900s is still used by some states today, Colorado now utilizes a standards-based approach. The history of teacher evaluations is also linked to the history of teacher unions, for teachers desired stability in their job beyond the whims of their principals (Spring, 1997). Indeed, the evaluations which led to teacher tenure were originally created, at least in part, to protect teachers, particularly female teachers, from the whims of their male principals. Although the methods of evaluating teachers have morphed over time, with some administrator textbooks even claiming that the focuses of supervision and evaluation shifted in the late 1900s from summative to formative assessments

(Sergiovanni & Strarratt, 2002), the current concerns about the value of teacher evaluations illustrate that this issue remains primary.

Like much of public education, teacher evaluations are a state issue; consequently, the types of teacher evaluations being used differ from state to state. In fact, depending on the state, some teacher evaluations differ significantly from school district to school district within a state (Weisberg et al., 2009). Some districts use a binary system where teachers receive either a "satisfactory" or "unsatisfactory" rating. Others, such as those in Colorado, have a standards-based system, though the enforcement of rigor differs among districts. Regardless of which approach to evaluations is used, though, critiques abound.

The standards-based system is more rigorous than the binary teacher evaluation systems but the standards-based system is highly subjective (Kimball & Milanowski, 2009; Gutherie, 2005; Alicias, 2005; Weisberg et al., 2009). One study noted that "evaluators' decisions were found to be a complex and idiosyncratic function of motivation, skill, and context" (Kimball & Milanowski, 2009, p. 34). In this particular study, evaluators were aware that their evaluation methods were under the magnifying glass and, thus, one might anticipate that such evaluators would be particularly careful to be objective, yet the study's authors still concluded that the evaluation process was significantly influenced by the mindset of and pressures facing the evaluator.

General concerns about the effectiveness of teacher evaluations, though, are not new. Challenges to the teacher evaluation system go back at least as far as the late 1800s

(Danzer, 2003). A sharp critique of teacher evaluations was written in the New York Times in 1936:

> There are at least 'several hundred' incompetents now in the school system [says the superintendent]. Other observers think there are several thousand, while still others insist that 'several' would be nearer the mark. Whether these incompetents were unfit to teach at any time, or have been rendered unfit by the passing years, is a matter of opinion. The question is, why are they allowed to remain? (Bernstein, 1936; found in Weisberg et al., 2009).

Bernstein's question, why "incompetent" teachers are allowed to remain in the classroom and are not weeded out through a viable teacher evaluation system, is one of the driving concerns behind the efforts to reform the teacher evaluation system today.

A review of current teacher evaluations in the United States reveals a system focused on functionality and not effectiveness. Most teacher evaluations across the nation make few distinctions among teachers. The New Teacher Project's (Weisberg et al., 2009) scathing report on teacher evaluations, entitled "The Widget Effect," found that, for those school districts in the study that used a binary evaluation rating, with teachers being either "satisfactory" or "unsatisfactory," 99% of teachers received a satisfactory rating. Moreover, the study found that regardless of which type of evaluation was used, less than 1% of teachers received "unsatisfactory" ratings (p. 6). In line with this finding, a Denver Post article (Meyer, 2010a) reported that only 32, or 1.3%, of Denver Public Schools' 2,387 non-probationary (i.e., tenured) teachers received an "unsatisfactory" rating over a three year study.

The overwhelming "satisfactory" rating that teachers' receive is at odds with national and international data about student achievement. In *Education Week's* 2013

annual "Quality Counts" report, Colorado received a "C" grade for the quality of its education, which fell just below the average state grade of a C+ (*Education Week*, 2013). In 2009, of the 34 OECD countries participating in the Programme for International Student Assessment (PISA) assessment, USA's students ranked 14th in reading, 17th in science, and 25th in math (OECD, 2011, p. 26). As these individual pieces of data collectively illustrated, America's schools and students are struggling. While there are many causes of the gap between teachers' strong satisfactory ratings and students' low international rankings, the teacher evaluation system may be a contributing factor. "The Widget Effect" report concluded that school districts in general, and the teacher evaluation system in particular, treat teachers as widgets, items that can be easily interchanged. The report asserted that "pervasive and longstanding failure to recognize and respond to variations in the effectiveness of our teachers" has led to a system that assumes "classroom effectiveness is the same from teacher to teacher" (Weisberg et al., 2009, p. 4).

New Methods of Teacher Evaluation in the Age of Accountability. In recent years, most likely as a result of both the national focus on testing and the data highlighting the importance of teachers in the classroom, there has been a renewed interest in how best to evaluate teachers. Indeed, efforts to determine precisely how valuable individual teachers are to their students' learning have sparked a new method of assessing teachers: value-added evaluations. First developed in Tennessee by William Sanders in the late 1990s, the Tennessee Value-Added Assessment System, which was later expanded and renamed the Education Value-Added Assessment System (EVAAS),

employs a mixed-model statistical methodology that tracks vast quantities of student achievement data over time for a school district (Alicias, 2005; Amrein-Beardsley, 2008). Statisticians use complicated algorithms to analyze multiple data points to determine whether an individual teacher has added value to his/her students over the course of the school year. The popularity of the EVAAS model seems to derive from its claims of not being influenced by students' backgrounds, not being "compromised" by missing student data, and for being able to handle massive volumes of data (Amrein-Beardsley, 2008, p. 66). Many of these same elements have led to profound criticism of the model.

The critics of the value-added approach generally fall into two (often overlapping) categories: those who are concerned about psychometric properties and statistical elements of the value-added approach, and those who oppose the use of student assessment data to evaluate teachers in general. Since EVAAS was developed and nationally recognized, others have been working to create algorithms to ascertain a teacher's effectiveness (McCaffrey et al., 2004; Martineau, 2006; Mariano, McCaffrey, Lockwood, 2010). Through the use of various forms of value-added models, each study concludes that student achievement data may technically be used as a method of assessing teacher effectiveness but it must be done with caution. There is the possibility of model misspecification, originating from such sources as missing student data, and this could have negative impacts on resulting teacher evaluations (McCaffrey et al., 2004). Another concern is that the value-added models readily experience construct shifts that render the data easily misunderstood (Martineau, 2006). In other words, value-added models are used for assessing the impact of a teacher in a variety of ways, and the

construct being tested (i.e., the academic idea, skill, fact, etc.) may not always be perfectly accessible by a particular test, whereby the data produced are more a result of the assessments' effectiveness rather than a true measure of a teacher's effectiveness. Educator effectiveness is supposed to be assessed from student achievement growth results, but the quantity of data needed to make such assessments is so broad, in terms of grade levels and content, that the data results are prone to shifting constructs that render the final estimates of teacher effectiveness "unacceptable for high-stakes use" (Martineau, 2006, p. 57). Even the researchers behind the cutting-edge value-added models acknowledge that their model is not the answer, but merely another step along the way to being able to accurately use student achievement growth data to assess teacher effectiveness (Mariano, McCaffrey, Lockwood, 2010). Indeed, each researcher or group of researchers of these new algorithms concludes his paper by raising significant concerns about the application and utilization of these models at this stage, given the serious consequences the results may carry for students and teachers.

In addition to concerns about the actual scientific properties of the value-added approach, there are some who challenge the use of value-added measures because they do not believe student data should be appropriated in this way to evaluate teachers. In Popham's *Assessment for Educational Leaders* (2006), he contended that both traditional standardized achievement tests and standards-based accountability tests are "instructionally insensitive" (p. 355). Achievement tests, he stressed, should not be used to assess student knowledge because there is often a mismatch between what is being taught and what is being tested, there are psychometric tendencies to "remove items

covering the most important things teachers teach" (p. 355), and the need for a score spread often makes it difficult to assess what knowledge teachers gave and what students brought to school. The concerns with standards-based tests include that tests try to assess students' mastery of "too many items," the states that design the tests do not give adequate information to teachers so that teachers may prepare students appropriately for the assessment, and the assessment results come out so late in the school year that they are useless for improving or altering instruction (p. 355). In total, Popham asserted that "[t]o evaluate educators' instructional effectiveness using assessment tools that deliberately avoid important, teacher-taught content is fundamentally foolish" (p. 338).

The counter-argument to those who generally oppose the value-added approach is that this new system protects students instead of teachers and, in fact, prioritizing students over teachers is the right approach (Glazerman et al., 2010). The argument is that value-added models are perceived as threatening to teachers because no teacher wants to be mislabeled as a bad teacher, because "no rate of miscalculation is acceptable" for the value-added models (Glazerman et al., 2010, p. 6), yet the current evaluation system enables the proliferation of poor teachers who hurt students' academic opportunities. Proponents of the value-added approach counter that it is high time teachers were held to a higher standard of evaluation for the sake of children. Moreover, while far from a perfect formula, if value-added models provide additional information that may enrich conversations about and decisions regarding personnel, then they are useful (Briggs & Domingue, 2011) and should be included in the evaluation process

(Glazerman et al., 2010). In these situations, the proponents of value-added evaluations argue that having something is better than nothing.

A less prevalent alternative to the value-added approach is the growth model, which Colorado itself adopted. A Growth Model (GM) is a statistical model used to calculate students' performances on statewide assessments. Results from the GM can be used for student, school, and district analyses (McGrane, 2009). GMs fill a special niche in testing, for they can be used to determine growth even if students are considered low performing (U. S. Department of Education, 2011). In this way, the GM is primarily used as a comparative tool. Colorado was not the first state to utilize GMs, but they have been in Colorado since 2008. In Colorado, the GM has been used to illustrate if a student, school and/or district is performing the same as, better than, or worse than other students with similar Colorado State Assessment Program (CSAP) data from years past. Many of the GM results come in the form of percentiles, such as student growth percentiles, which focus on overall growth from the student's starting point rather than just focusing on the test results from that particular year. McGrane (2009) stated that a low performing student demonstrating 90^{th} percentile growth is more impressive than a high performing student demonstrating 10^{th} percentile growth (p. 10). Although the GM offers the chance to observe student, school and/or district growth, the GM model is often critiqued because different GM models produce different results with the same data (U.S. Department of Education, 2011, p. 98).

Both the value-added model and the growth model have developed in recent years, at least in part, because of the aforementioned increased national focus on testing

and accountability. The testing of American students began in the early 1900s as an off-shoot of the testing developed for soldiers entering World War I, but was primarily limited to the administration of Stanford-Binet (i.e., IQ) tests to students in wealthier school districts (Graham, 2005). Annual testing of students began in the 1960s, though at this time it was largely limited to students in schools serving high poverty communities (Forte, 2010). Testing practices changed significantly under the Bush Administration's NCLB Act of 2001. Although NCLB received bipartisan support at the time of its passage, it has come under significant critique from multiple angles over the years.

One of the primary objectives of NCLB was to reduce the achievement gap between "have" and "have not" students, specifically between white students and students of color (Torres, 2004), by improving the achievement of the "have not" students. One of the primary means to do this was through the annual testing of large groups of students to determine the progress of these groups of students and schools overall. It is important to note that the focus on testing as a method to reduce the achievement gap has itself come under serious scrutiny. The increase in high-stakes testing has led some teachers to abandon their traditional curriculum for teaching to the test (Ravitch, 2010), and this is particularly true in schools serving large percentages of low socioeconomic status (SES) students (Stone & Lane, 2003). At this time, is unclear if teaching to the test helps students (Nichols, Glass, & Berliner, 2006), and doing so may actually hurt students' overall academic achievement (Pedulla et al., 2003; Berliner, 2005a) and well-being (Noddings, 2001). It is possible that the law designed to minimize

the achievement gap may actually exacerbate it due to the high-stakes results associated with the increased testing processes.

Educational historian Diane Ravitch (2010) moved from being a prominent supporter of NCLB, with its focus on testing and accountability, to strongly opposing it precisely because she believes that the focus on testing and accountability has led to *decreased* learning. Learning has multiple definitions, ranging from content mastery to the honing of specific skills, yet the seemingly sole focus in the age of accountability is content mastery. Almost a century ago, John Dewey expressed concern that students were not learning to think in school (Dewey, 1916). Thinking, Dewey contended, requires more than exposure to a topic or idea; thinking, and the training of thinking, requires an individual "wrestling with the conditions of the problem at first hand, [and] seeking and finding his own way out" (Dewey, 1916, p. 152, 160). Dewey was not worried about a national obsession with testing, only that the material presented in school required adequate thinking. Individuals who know how to take a test, but who do not know how to analyze ideas and make meaning of content, may be able to score high marks on a test but it is unclear whether they are prepared to participate fully in society or to support a well-functioning democracy. The entire conversation about character education (Noddings, 2005; Graham, 2005) and education for participation in a democracy (Mann, 1872), is secondary during this age of accountability, as is the purpose of public education in general.[2]

[2] This is not to imply that conversations about the purpose of public education are absent (Noddings, 2005; Graham, 2005). In the current era of accountability and testing, the national focus is on students learning to be able to obtain jobs and contribute to the economic health of the country. Concerns about character

Despite these challenges to NCLB itself, the testing and accountability movement remains strong. In 2009, the United States Department of Education sponsored the Race to the Top Fund (RTTT). Launched under President Obama's administration, RTTT provided a competitive format in which educational leaders in the 50 states could compete for much-needed financial support for their state's public education system. With a total of $4.35 billion to be distributed to winning states, RTTT aimed to "spur systemic reform and embrace innovative approaches to teaching and learning in America's schools" (White House Press Release, 11/4/2009). Rules for applying to RTTT were specific and extensive. The RTTT application had numerous requirements, including Requirement D: Great Teachers and Leaders. Requirement D(2) called for measuring student growth and connecting this student growth with teachers and principals through annual evaluations (U. S. Department of Education Executive Summary, 2009).

Colorado's SB 191 developed during the RTTT application phase. The full name of SB 191, "The Great Teachers and Leaders Act," let alone the details within the bill, connects directly to RTTT specifications. Of the six selection criteria detailed in RTTT's executive summary, it was the fourth section or criteria "D: Greater Teachers and Leaders" that links to SB 191. Specifically, D(2) outlined how to improve "teacher and principal effectiveness based on performance" (U.S. Department of Education, 2009), with multiple subcategories requiring the measurement of student growth; the implementation of "rigorous, transparent, and fair evaluation systems for teachers and

---

development, particularly "the secular traits of integrity, ingenuity, and hard work, both individually and collectively, that our democracy needs" appear to be on the backburner (Graham, 2005, p. 253).

principals" that use multiple categories, including student growth; and the annual evaluation of teachers and principals. SB 191's requirement of annual evaluations is emblematic of the RTTT, for previously Colorado's probationary teachers were evaluated annually, but once a teacher achieved non-probationary status after three years, he/she would be evaluated only every five years (unless something significant caused this to change). Additionally, prior to SB 191, evaluations were standards-based and student growth had no direct bearing on teacher or principal evaluations.

SB 191 is situated within the testing and accountability movement, with parts of the bill specifically focused on the teacher effectiveness and evaluation reforms. SB 191 is the newest in a long line of educational reform efforts. A review of the USA's history of modern education reforms reveals efforts gone awry.

Brief History of and Challenges to Modern Education Reform

The purposes of public education have changed over the centuries and are still evolving. The law that laid the groundwork for the USA's public education system, often called the Old Deluder Satan Act, required all Massachusetts townships to maintain a school and hire a teacher so that children would learn to read the Bible and ward off the Devil (Hlebowitsh & Tellez, 1997). In the 1830s, the Common School Revival, led by Horace Mann, sought to educate all students so that they would become politically aware, active citizens and have the character necessary to live in and contribute to a democracy, both of which would reduce societal stratification. The current focus is content mastery and international competition in a global economy, with secondary concerns about caring for and educating the whole child (Noddings, 2005; Graham, 2005).

As the purpose of public education continually evolves, so do the reform movements. Graham's (2005) recent history of the USA's public education system noted that the 20th century experienced four large waves of reforms: assimilation; adjustment; access; achievement and accountability. Each wave brought some substantive changes. The public education system of the 21st century is significantly better than it was 100 years ago; progress is being made. Yet, Graham notes that "educational practice changes slowly" and that the reforms ultimately instituted are "usually only a partial implementation of the new idea" (Graham, 2005, p. 249).

While many reforms failed to make significant inroads, some reforms become so ingrained as to be considered part of the "grammar" of schooling (Tyack & Tobin, 1993, p. 454). Tyack and Tobin (1993) wrote that two elements considered nearly essential in public schools are the graded classroom and Carnegie units of study, yet these concepts and practices developed out of practical necessity during the Industrial Age and may not serve students well in the 21st century (Robinson, 2010; Senge et al., 2000). They went on to note that previous efforts to undo these measures, such as the Dalton Plan and the Eight-Year Study, failed to gain long-term traction and ultimately disappeared from practice. The question of why certain reforms hold and others fade, argued Tyack and Tobin, rests with two factors: the failed reforms were "too intramural," meaning that the proponents of the reform did not have the social capital to sway others to adhere to the new ideas, and the failed reforms engendered "turnover and burnout" from the reformers (pp. 477-478). Certain reform measures collapse not because they are poorly conceived or inappropriate for schools, teachers, and students, but because the messages behind

them gain insufficient traction and/or they require too much energy beyond what was being already required for those implementing them.

This finding is consistent with others' work (Sarason, 1990; Cuban, 1998; Tyack & Cuban, 1995). Tyack and Cuban supplemented this conclusion in their award-winning work *Tinkering Toward Utopia* (1995), finding reforms that try to alter foundational elements of public education run amok when reformers underestimate the "stickiness" of traditional school structures and when the reform movement overloads practitioners (p. 108). This stickiness factor may simply be educational habit (Graham, 2005), but it can also be more meaningful. Sarason (1990) noted that top-down reforms, such as those created at the state legislature that then filter down to districts and schools, can be legitimately challenging for teachers and building administrators to stomach because such reforms often upend years of practice and perhaps even the very pedagogy that introduced said practitioners into the practice of teaching. Educational reformers, Sarason contended, "seem unable to understand what is involved in *unlearning* what custom, tradition, and even research have told educational personnel is right, natural, and proper" (Sarason, 1990, p. 101; italics added). Additionally, the overload factor occurs frequently because "[r]eforms have rarely replaced what is there; more commonly, they have added complexity" (Tyack & Cuban, 1995, p. 83). This pattern can lead some teachers to "drag their feet" during the implementation phase, hoping to wait out a particular reform by "reassuring themselves that this, too, shall pass" (Tyack & Cuban, 1995, p. 135). Thus, teachers and administrators may be reticent to comply with new reforms because of ideological differences, but practitioners may also challenge—and even thwart—

implementation simply because the new reform is one change too many for them to incorporate along with everything else.

Yet another challenge to implementing reforms with fidelity is the time pressure that often accompanies reforms. Currently, many reforms have strict timetables written into the law and severe consequences for those schools/districts/states that are unable to fulfill them. A front page story in the Denver Post in November 2012 (Auge) noted that five schools failing to make mandated changes in a short time period were going to lose significant funding in the upcoming school year. The timeline for Colorado's SB 191 was written into the bill itself, with a clear mandate to have the entire state using the newly created teacher and principal evaluations for the 2014-15 school year. Strict timelines can present a serious challenge to reforms, for creating sustainable change takes time (Hargreaves & Fink, 2006). Yet, it is improbable to create a reform without some sort of time schedule, for otherwise there would be no structure compelling individuals to enact the reform.

In addition to the aforementioned challenges, there is a level of distrust between those who write the law and design implementation and those who are responsible for actualizing the new policies in schools with students. As Larry Cuban noted, "schools change reforms as much as reforms change schools" (1998, p. 455). Yet the opinions of teachers and administrators, the ground-level actors of reform efforts, are frequently discounted (Sarason, 1990). In recent years, there have been a few cases that gained national attention when school boards fired all teachers in a school or even a district as the starting point of systemic changes to improve the school or district (Zezima, 2010;

Goodnough, 2011). The act of terminating teachers to solve a school's academic failures implies that the fault for such failure lies with teachers and administrators. Rather than working with teachers to solve problems, teachers themselves are deemed the problem.

Underestimating practitioners is dangerous on many levels. First, the lack of trust between policymakers and practitioners is detrimental in and of itself (Senge et al., 2000; Sarason, 1990). Currently, Finland is seen as the gold standard for education practices in 2012 (Sahlberg, 2010) and one of the key reasons cited is the trust in teachers (Sahlberg, 2010; OECD, 2011). Second, distrust of practitioners is shortsighted and misplaced, for in practice teachers are "street level bureaucrats" who are empowered with "sufficient direction, once the doors close, to make decisions about pupils that add up over time to de facto policies" regardless of the official policies (Tyack & Cuban, 1995, p. 135). In practice, some teachers may fail to enforce a new policy in such a way that it ultimately undermines the entire reform; other times, teachers may amend the reform to make it more applicable to schools and students (Cuban, 1998). This second approach raises issues of fidelity, but it also recognizes that ground-level actors may have the ability to design and actualize a modified reform to reach the same desired outcome as the original reform.

The frequent inability of reforms to create the intended changes for myriad reasons is highly problematic (Sarason, 1990). Current reform efforts appear to be falling into this pattern. Despite 2001 NCLB's mandated reforms to close the achievement gap between particular student groups, those gaps still exist in 2013. So, too, does the USA's annual poor performance on and rating in international assessments like PISA. The

nation's current concentration on teacher effectiveness and evaluations, within the testing and accountability movements, will likely remain a sharp focus of efforts for years to come. If history reveals patterns, though, then chances are poor that any long-term, significant changes will develop as a result of these efforts.

But patterns can shatter. Although there is no single solution, the realization that teachers and administrators may directly contribute to the failed reform efforts, or alternatively, their success, merits a closer look. In 1990, Sarason recommended the paradigm shift of recognizing that reforms impact both students *and* teachers. Schools are filled with teachers and building administrators who need time to think and adapt to new systems (Senge et al., 2000; Hargreaves & Fink, 1996; Hall & Hord, 1987). Schools themselves are systems and systems take time to change, at least to change authentically and sustainably (Senge et al., 2000; Hargreaves & Fink, 1996). Thus, in order to get to the heart of reforms, one must understand the organization of schools themselves first. If long-term, systemic changes are going to occur throughout public schools, then these reforms must incorporate the needs of students *and* teachers and be designed for sustainability.

Organizational Change and Sustainable Leadership

Organizational change and sustainable leadership deserve exploration in an effort to understand how schools, the people in them as well as the organizations themselves, can best adapt to new changes in a meaningful and sustainable way.

According to Bolman and Deal's *Reframing Organizations: Artistry, Choice, and Leadership* (2008), organizations function through four frames: structural, political,

human resources, and symbolic. Bolman and Deal contend that in order to make lasting organizational change, one must consider how a particular change impacts all four areas and not just one or two areas. If organizations are as multifaceted as implied, then implementing reforms with fidelity is that much more complicated.

Effective organizational change requires consideration of the people directly involved in the change process. Wheatley (1999) succinctly noted, "[w]e live in a universe where relationships are primary" (p. 69). Every person may approach organizational change from one of the four frames, but there are also personal factors to consider in relation to people's readiness to adopt a reform or change. How teachers and school leaders perceive a reform impacts how the reform is implemented (Hall & Hord, 1987, p.5). Often workshops, professional development, and the like that are designed to improve the craft of teaching are designed by the facilitators rather than being driven by teachers' own needs (Hall & Hord, 1987); school improvements might have a greater chance of success if teachers' needs were considered directly when designing such programs (Hall & Hord, 1987; Sarason, 1990). Certain methods to school improvement, such as Hall and Hord's (1987) Concerns-Based approach, are grounded in the assumption, among others, that "understanding the point of view of the participants in the change process is critical" (p.8). This concept is important not only because individuals are often more invested in making changes when they believe in the changes, but also because making changes takes significant effort. In many ways, a teacher's approach to teaching can become habitual. How a teacher approaches the subject matter, designs lessons, composes assessments, and the like are often patterns. As Duhigg's (2012) work

The Power of Habit illuminated, habits can become addictive and it takes deliberate, constant effort to change them. Thus, asking individuals to take a significantly different approach to teaching and to their job status, as SB 191 will do for many, may create significant disruption in practitioners' lives. It is possible for teachers to break habits, adjust to the change, and perhaps improve their craft in the process. But changing habits takes time and practice (Duhigg, 2012). It is personal (Hall & Hord, 1987, p. 330). The same can apply to institutions that possess routinized approaches.

Hargreaves and Fink (2006) identified seven pillars of sustainable leadership: depth, length, breadth, justice, diversity, resourcefulness, and conservation (p. 18). They argue that leadership that does not account for sustainability is short-sighted and prone to weakness. In order for a reform to have the potential for long-term success, building educators and policymakers must develop and maintain sustainable leadership. It is insufficient for leaders to institute a change in name; practices must be changed and all impacted participants convinced that the change will be beneficial to the school, students, and teachers in the long run.

The goal of improving student learning through the revamping of the teacher evaluation process, as in SB 191, requires a paradigm shift for individuals working in school buildings, school districts, and the state capital about the value of teaching and the standards for teachers. Sustainable leadership is needed at all levels of the educational system to enable the successful implementation of SB 191.

Exploring organizational change and sustainable leadership is critical to understanding how schools can best incorporate significant policy changes into their

practices in a meaningful way. The theoretical roots underlying tension between teaching and teacher evaluations, testing and learning, top-down policy changes and sustainable reforms, can be illuminated through a critical social theory lens.

Critical Social Theory and Change

Critical social theory has many branches and can be challenging to codify succinctly. Levinson et al. (2011), in *Beyond Critique*, defines critical social theory as "those conceptual accounts of the social world that attempt to understand and explain the causes of structural **domination and inequality** in order to facilitate human emancipation and equity" (p. 221; bold in original). Critical social theories examine societal assumptions and seek to make explicit what is often implicit (Levinson et al., 2011). Critical social theory is applicable to public education in the USA in the 21st century because public education exerts a great influence on society. In fact, the American Dream espouses that, regardless of a student's background, if a student works hard in school and excels, then that student will have the chance to succeed (implicit is that this success will be financial). Critical social theorists challenge this assumption, arguing that the opposite is often true: schools reproduce the status quo (Levinson et al., 2011). Rather than creating a system designed to give all an equal opportunity for success in life, the public education system provides a way to maintain the existing power relations. "[E]ducation is a big part of the way social structures do their work to distribute power and knowledge and life chances unevenly" (Levinson et al., 2011, p. 15).

This study was rooted in critical social theory, with a focus on theories of power and voice as expressed through Gramsci, Bourdieu, and Foucault. Gramsci explored

critical consciousness and its challenging of the hegemonic order, with critical consciousness defined as an individual's ability to reflect "critically on one's position in society, relative to broader social structures" and hegemonic order being the nuanced ruling order of the society (Levinson et al., 2011, p. 56, p. 58). Applied to public education, Gramsci argued that the teachings of public education were determined by society's elite and, therefore, deliberately masked the existing hegemony which curbed students' abilities to develop their critical consciousness. The elite held the power to determine what was taught and measured, thereby creating an environment where the status quo could continue without challenge. Bourdieu picked up on this structural inequity in the school system and expanded it, going so far as to pronounce the system one of "symbolic violence" (Levinson et al., 2011, p. 123). Bourdieu believed much of education, such as curriculum selection, assessment measures, and even teachers' interactions with students, was "arbitrary" and reflected the elite's "cultural capital" over the masses rather than a true understanding of knowledge or best practices (Levinson et al., 2011, p. 123). The symbolic violence is the deliberate reproduction of stratified social classes hidden behind the appearance of equality for all. The voices of the masses are secondary to those of the elite under the false pretense that all are regarded equally and have equal opportunities to excel in academics and society.

Foucault explored power from a different lens, namely one that understood power to be a "multiplicity of force relations" rather than something located in an individual or group (Levinson et al., 2011, p. 144). Applied to public education in the USA, power can be seen as the vocabulary and phrases preferenced by those with positions of authority,

the type of learning most valued, and the measures of knowledge most revered. In other words, power is not simply who is in a position of authority, but how those in positions of authority use their titles to set national standards about what learning is, based on such things as content standards for each subject, types of assessments, and measures of effectiveness.

As the theoretical framework for this study, critical social theory—specifically the issues of voice and power—was applicable in two strands. The first applies to concepts underlying SB 191 and the second to the research, namely the desire to learn of teachers', administrators', and policymakers' perceptions of their roles in the implementation of this new policy.

From a critical social theory perspective, the current testing and accountability movement, in which SB 191 is rooted, raises issues of power and voice. Who determines what it means to learn as well as how to assess knowledge dictates the purpose of public education in the 21st century. Policymakers have declared annual statewide content tests for all students in most grades as the best way to assess knowledge. The decision to use formative assessments in this matter is also a decision to allow a postpositivist perspective to dominate education. In the wake of positivism's death (Phillips, 2004), postpositivism thrives. Postpositivism posits there is a singular reality that can be studied through deductive reasoning and that researchers will be distant, impartial, and unbiased (Creswell & Plano Clark, 2011). There is one answer to a problem and one can learn all one needs to know from a single platform, such as a formal assessment. The postpositivist perspective is embodied by the various state and federal legislations that

place greater and greater weight on test scores to determine individual student learning and overall school progress—and now teacher effectiveness, too.

A postpositivist approach to education stands in stark contrast to those who view education as more than the sum of assessments, but as the opportunity to educate a whole child (Dewey, 1916; Noddings, 2005). The choice to focus on multiple elements of educating a child is best embodied by the constructivist approach. Constructivism veers further from positivism than postpositivism by allowing for multiple constructed realities and by welcoming inductive research that comes from researchers who are close to their topic and/or subjects and acknowledge that their biases influence how they perceive results (Creswell & Plano Clark, 2011, p. 42). Applied to public education, a constructivist recognizes there are multiple answers to most questions and there is no single solution to large problems.

The tension between the postpositivist and constructivist approaches to public education are palpable in schools, but often swept aside in public policy because of power and voice. With SB 191, how students perform on a test will be a greater measure of teacher effectiveness than whether a teacher inspires learning, is able to connect with students and keep them engaged, or is able to support the child's overall development. Although SB 191's teacher evaluation system uses factors other than student growth data results, those other factors are collectively weighted at 50% while the assessments are also 50%. The harder-to-measure elements that address the whole child are minimized, despite how essential they may be to students' well-being.

From a critical social theory perspective, issues of power and voice are also at work in the implementation of reforms such as SB 191's new teacher evaluation system. Those who wrote and passed SB 191 are policymakers, politicians who hold substantial social capitol in American society; those who are responsible for implementing the bill are educators, a group whose members generally do not hold substantial social capital. This imbalance in social capital, an imbalance of power between who makes the laws and who is responsible for implementing them, is real. In many ways, the imbalance in power between who creates laws and who implements reforms mimics the public school system's reproduction of stratified social classes that Bourdieu noted (Levinson et al, 2011, p. 119). Additionally, the issue of voice—who is able to determine which components are problems in the educational system, what needs to be fixed promptly, and what solutions are appropriate—arises when considering the implementation of SB 191. Colorado policymakers determined that Colorado's students' learning was inadequate, that changing teacher and administrator evaluations was the way to fix it, and then outlined the methods to change them as well as the timeline for doing so. Although the lead author of SB 191 was a former teacher and administrator himself, he was able to use his social and cultural capital as a legislator to write a bill, gain support for it, and see it through. The collective thoughts and concerns of current practitioners were consulted marginally (Meyer, 2010b), but now these individuals are responsible for enacting the reforms. As Sarason (1990) noted, teachers and administrators are integral parts of actualized reforms. In taking lessons from organizational change, it is imperative to have those responsible for enacting changes to have a voice in the changes. One of the central

aims of this study is to give voice to teachers and administrators in the Rockies School District, seeking to hear those who often go unheard in education reform policy matters.

There is yet another imbalance of power regarding the implementation of SB 191: the different power structures within schools, for administrators have power over teachers. With the elimination of teacher tenure as part of SB 191, each teacher evaluation holds greater significance for a teacher's job security and, therefore, the imbalance of power between administrators and teachers increases. This power structure could impact teachers' perceptions of their own voice within their schools, as speaking one's thoughts could have higher consequences in the future.

From a critical social theory perspective, the question of why practitioners' voices are not invited into the dialog about policy changes is easy to answer: honoring practitioners' voices would disrupt the existing power structure, granting power to a group that regularly has little. Policymakers, who wield greater power than in-building educators, would have to be willing to relinquish some of their power and authority in order to give voice to education practitioners. To challenge the critical social theory perspective, though, it is untenable to expect that all teachers and administrators could be regularly involved in the development, design, and implementation of statewide education reforms. Teachers and administrators are busy teaching, grading, planning, and leading in their own schools. It is reasonable to have another group of professionals devoted to thinking about statewide policy on a regular basis.

This study utilizes the critical social theory approach as it seeks to give voice to those often perceived as powerless or less powerful. The theoretical framework provides

insight into how schools might better incorporate state-mandated changes. Stakeholders' perceptions of SB 191 are connected to how stakeholders understand the purpose of education and their own role in actualizing that purpose for their students and school.

Summary

In this era of testing and accountability, significant focus is on teacher effectiveness and teacher evaluations. Colorado led the way with SB 191, but many states now have similar laws. Given the sad state of recent reform efforts to make sustainable improvements in student learning on a national scale, the odds are against SB 191 improving student learning by overhauling the teacher and principal evaluation systems and eliminating teacher tenure. Yet the overview of organizational change and sustainable leadership illustrated that authentic, sustainable change is possible, if schools' teachers and administrators are involved in the change process. Adding teachers' and administrators' voices into the implementation design may be a key factor to the long-term success of a statewide mandate.

Chapter 3: Methodology

Introduction

The objective of this mixed methods study was to explore how teachers, administrators, and policymakers perceived their role in the implementation of SB 191. This chapter describes the research questions and study design, site selection and population, instrumentation, data collection, confidentiality, data analysis, limitations, and researcher bias.

Research Questions and Study Design

The research questions for this study were:

1) How do Colorado's teachers, administrators, and policymakers perceive their respective roles in the implementation of SB 191's new teacher evaluations?

2) Do Colorado's teachers', administrators', and policymakers' responses vary by demographic variables?

3) How do the responses of Colorado's teachers, administrators, and policymakers compare?

To obtain information about how the stakeholders perceived their role in the implementation of SB 191, the researcher used an explanatory sequential design (Creswell & Plano Clark, 2011). This design has two distinct yet connected phases: a quantitative phase followed by a qualitative phase that is designed to elaborate upon interpretation of the quantitative results. In this mixed methods approach, the "second,

qualitative phase of the study is designed so that it follows from the results of the first, quantitative phase" (Creswell & Plano Clark, 2011, p. 71). This study began with the development and pilot testing of an online survey for the stakeholder groups of interest, proceeded to the dissemination and data analysis of the online survey for the quantitative phase, and concluded with a community-based research (CBR) approach to interviews and the constant comparative analysis of interview data.

For over three decades, a series of methodological "wars" existed in the social and behavioral sciences "regarding the superiority of one or the other of the two major social science paradigms" (Tashakkori & Teddlie, 1998, p. 1). In the social and behavioral sciences, the methodologies of the positivist, empiricist approach battled the constructive, phenomenological approach with a resolution occurring in the mid-1990s due to the development of a pragmatist approach (Tashakkori & Teddlie, 1998). For years, researchers debated if postpositivist approaches, generally quantitative designs, were compatible when united in a single research study with constructive approaches, generally qualitative designs.

The primary concern with the union of these two methodological approaches was that the different underlying philosophies of each approach made the combination of their designs inherently contradictory. For instance, a postpositivist approach takes the ontological perspective that there is a single reality and then uses quantitative data to provide the single answer to the research question. In a constructivist's ontology there are multiple realities, as illustrated by the varied nature of, for instance, each interviewee's response to a single question. Experts who questioned the combination of method designs

in a single study asked how one can reasonably reconcile these contradictory approaches to the nature of reality, let alone all other philosophical concepts. The answer was the development of the pragmatic design model that is mixed methodology (Tashakkori & Teddlie, 1998).

Like postpositivist and constructivist paradigms, the pragmatic model contains specific approaches to the worldview elements of ontology, epistemology, axiology, and methodology. In pragmatism, the nature of reality is defined as being both singular and multiple, meaning that the value of both approaches is recognized. Pragmatism holds a practicality approach to epistemology (Creswell & Plano Clark, 2011), meaning that researchers are expected to use whatever means are most appropriate for gathering data for the particular subject, rather than being limited solely to quantitative or qualitative approaches. The axiology of pragmatism is the reality that researchers are both biased and unbiased in their collection and interpretation of data. The methodology embodied in pragmatism is that of combining; quantitative and qualitative data should both be collected and mixed to produce the best possible data and analysis results.

With the closing of the methodological wars in the 1990s (Tashakkori & Teddlie, 1998), it is acceptable, often preferable in education fields, to utilize a mixed methods design (Greene, Caracelli, & Graham, 1989). One of the benefits of the mixed methods design is its reliance on triangulated data sources. Triangulation is loosely defined as the use of multiple methodological approaches to study the same problem (Jick, 1979). One of the critical ideas behind triangulation is that it compensates for the natural weaknesses of each method. As Green, Caracelli, and Graham noted:

> [A]ll methods have inherent biases and limitations, so [the] use of only one method to assess a given phenomenon will inevitably yield biased and limited results. However, when two or more methods that have offsetting biases are used to assess a given phenomenon, and the results of those methods converge or corroborate one another, then the validity of inquiry findings in enhanced (1989, p. 256).

Triangulation can provide the best means by which to examine a single construct. Triangulation has occasionally been able to go beyond the mere merging of two approaches to "capture a more complete, *holistic*, and contextual portrayal" of the subject being studied (Jick, 1979, p. 603; italics in original). Thus triangulation can produce a whole that is greater than the sum of its two parts. In this study, the two methods being triangulated were surveys and interviews.

Within mixed methods designs, the explanatory sequential design is most appropriate when qualitative data are needed to explain quantitative results (Creswell & Plano Clark, 2011). This design allows one to "assess trends and relationships with quantitative data but also be able to explain the mechanism or reasons behind the resultant trends" (Creswell & Plano Clark, 2011, p. 82). This method is also appropriate to use when desiring to identify groups for qualitative data collection based on the results of quantitative data (Creswell & Plano Clark, 2011). In this study, both reasons were applicable. To understand how Colorado's teachers, administrators, and policymakers perceived their role in the implementation of SB 191's new teacher evaluation expectations, it was important to disseminate surveys that would reach many individuals in all three groups. A wide distribution allowed for the collection of beliefs about teacher evaluation systems, thoughts about education reform and change, and perceptions about

each group's role in the implementation of this particular reform. But these data could not explain the "how and why" of the results. Thus, the quantitative data were complemented by interviews to gather personal narratives expanding on survey themes and trends in the data results. The qualitative data were essential in capturing the authentic voices of practitioners to better understand how teachers, administrators, and policymakers each perceived their role in the implementation of the new teacher evaluations. The triangulation of the survey and interview data provided a more holistic insight into this question than either approach could have done individually.

Site Selection and Population

The setting for this study was a metropolitan school district in Colorado, with additional insights from state level policymakers. The Rockies School District (pseudonym) (RSD) was a convenience sample site (Merriam, 2009) selected because it met several criteria important to the study: It is one of the larger and more diverse school districts in Colorado, and its size allowed for a substantial number of teachers and administrators to be participants. RSD has more than 25,000 students and more than 3,000 teachers and administrators.

In addition to teachers and administrators in the RSD, members of the Colorado 2013 Legislative Assembly participated in the study and comprised the policymaker group.

Teachers. All RSD teachers had the opportunity to participate in the survey. In SB 191, a teacher is defined as: "a person who holds an alternative, initial or professional teacher license" (Senate Bill 10-191, 2010). Yet, in CDE publications related to SB 191,

there is an additional group of individuals who work in schools under the broad label "Other Licensed Professional" who will also be subject to a new evaluation system. The CDE defines Other Licensed Professionals as: "school audiologists, nurses, occupational therapists, physical therapists, psychologists, social workers, speech-language pathologists, counselors and teachers on special assignment" (Colorado Department of Education, 2012, p. 7). At the time of survey dissemination, it was known that these other licensed personnel were going to be evaluated by a new evaluation matrix, but unknown if members of this group would be evaluated using the new teacher evaluation system or something else. Consequently, to be both cautious and inclusive, all RSD employees who fell under the "Other Licensed Professional" category were invited to complete the survey for a total of 3,597 possible participants, under the broad heading of "teacher."

The list of individuals included under the general "teacher" heading from RSD included: classroom teachers, occupational therapists, speech language pathologists, deans, counselors, licensed coordinators, athletics directors, audiologists, mentors, specialists, psychologists, social workers, mental health coordinators, nurses, Licensed Practicing Nurses, and nurse mentors. This group list was generated through the RSD's human resource system and the list was then sorted by union code (teacher, administrators, etc.), location, and counts by union code. This list included active staff (no one on a Leave of Absence) and employees in their final year before retirement (R1-110), and excluded substitutes. The union codes used were: LICENSED (n = 3412), ADMIN (n = 102), MNTAL HL (n = 113), NURSE (n = 72).

Administrators. Participants were building administrators, defined as principals and assistant principals. In RSD, there were 102 administrators at the time the survey was disseminated. This list was generated using the process described above.

Policymakers. The policymaker participants were members of the Colorado General Assembly for the 2013 legislative session. The decision to limit the term "policymakers" to current members of the General Assembly was made for three reasons: 1) the inclusion of lobbyists, leaders of organizations that were heavily involved in the passage of SB 191, and state policymakers would have been cumbersome; 2) in the local media, state legislators were often pitted against educators (Haley, 2010; Meyer, 2010a; Lamm, Romer, Owens, & Ritter, 2010; Pena & Zellar, 2010); and 3) this was a convenient and succinct way to delimit this group. Members of the Colorado Legislative Assembly include individuals from both the Senate and the House of Representatives. The population of the 2013 Colorado Legislative Assembly was 100.

Survey participation. Out of the 3,799 individuals who may have received the email containing the link to complete the online survey, a total of 665 participants took the survey (Table 1). Although not everyone who started the survey completed every

Table 1
Response Rates

	Frequency	Total Possible	%
Teachers*	589	3597	16%
Administrators	64	102	63%
Policymakers	12	100	12%
Total	665	3799	18%

* Teachers includes classroom teachers plus everyone in the "Other Licensed Professional" category

question, the demographic data collected from the survey indicates that the teachers and administrators who did complete the survey were representative of the general RSD population (Table 2). There is a slight overrepresentation of female participants in the study, but otherwise it mirrors the district's population.

In addition to looking at the subgrouping of teachers and administrators by school level and gender, the subgroupings of teachers by teaching type (teacher of a classroom with TCAP assessments, teacher of a classroom without TCAP assessments, and "Other Licensed Professional" in a school), license type (Professional and Non-Professional) and ethnicity (minority and white) showed that the participating teachers came from a range of backgrounds (Table 3). For the ethnicity subgrouping, it should be noted that analysis for this was inadvertently complicated by the survey design. In an effort to make this question as open as is possible within a multiple choice framework, participants were able to select more than one option; however, the analysis software (SPSS), translated this option into six different questions, instead of six answers to one question. In order to best analyze these data, all those who checked an ethnicity option other than "white" were hand-coded into the "minority" category (as seen in Table 3)(to see the complete Ethnic breakdown of participants, see Appendix J).

Table 2

RSD Demographics and Survey Participant Demographics

		RSD Demographics				Survey Participant Demographics		
School level	RSD Label	Female	Male	Count		Female	Male	Count
Elementary		88%	12%	1673		88%	12%	257
	ADMIN	61%	39%	44		71%	29%	21
	LICENSED	88%	12%	1528		89%	11%	234
	Mtl HEALTH	89%	11%	57		100%	0%	1
	NURSE	100%	0%	54		100%	0%	1
Middle		71%	29%	786		76%	24%	127
	ADMIN	45%	55%	29		50%	50%	14
	LICENSED	71%	29%	730		80%	20%	113
	Mtl HEALTH	76%	24%	17		0%	0%	0
	NURSE	100%	0%	17		0%	0%	0
High		61%	39%	1065		60%	40%	156
	ADMIN	48%	52%	27		57%	43%	14
	LICENSED	61%	39%	1009		60%	40%	141
	Mtl HEALTH	83%	17%	18		0%	100%	1
	NURSE	100%	0%	11		0%	0%	0
District Wide								
	ADMIN	54%	46%	102		61%	39%	49
	LICENSED	76%	24%	3412		78%	22%	488
	Mtl HEALTH	83%	17%	113		50%	50%	2
	NURSE	100%	0%	72		100%	0%	1
TOTAL*		**76%**	**24%**	**3699**	**Sub-total**	**77%**	**23%**	**540**
					Count	*415*	*125*	
					Missing			113
					TOTAL**			**653**

* Includes Entire Population in Study

** Includes every survey participant who marked "teacher" or "administrator" in the first question, even if participant did not complete demographic information at end of survey

Table 3

RSD Teachers by Groupings (for 592 teachers total)

	Frequency	Percentage
TCAP	274	46%
Non-TCAP	161	27%
Other Educational Professional	144	24%
missing	13	2%
Professional License	477	81%
Non-Professional Licenses	40	7%
missing	75	12%
Minority	45	8%
White	468	79%
missing	76	13%

Interview participation. For the interview process, participants were selected through purposive sampling. Purposive sampling is defined as the deliberate selection of individuals for a study because they can "purposefully inform an understanding of the research problem and central phenomenon in the study" (Creswell, 2007, p. 125). Stated another way, purposive sampling allows the researcher to "discover, understand, and gain insight" into the research questions by selecting "a sample from which the most can be learned" (Merriam, 2009, p. 77).

Purposive sampling for this study was grounded in the community-based research (CBR) approach. CBR is defined as "collaborative, change-oriented research" that includes university members, often faculty and students, and community members (Strand et al., 2003, p. 5). It shifts a study from asking solely "how does one study *about*

a particular set of people?" to "how does one research *with* this particular set of people?" (Couto, 2003, p.70). In its purest form, CBR includes community members in the initial design of a research problem and question (Strand et al., 2003; Reason & Bradbury, 2001). However, this often is not feasible for many reasons (Strand et al., 2003; Couto, 2003), as it was not in this study. But research may still be considered to follow the CBR approach, provided that the study produces "information that meets CBR's most important criterion: usefulness to the community" (Strand et al., 2003, p. 10).

For this study, purposive sampling was combined with the CBR approach, and the researcher worked with one of RSD's assistant superintendents to select interviewees from both the teaching and administrative sides who would best be able to help interpret the initial quantitative data results. The assistant superintendent provided the names of 11 teachers and administrators in RSD who he deemed were exemplary in their fields and could add significant insights into the findings; ultimately, six individuals participated in the interview process. In keeping with the CBR tradition, these individuals were also able to provide recommendations for how the research should be used once the study was over (Strand et al., 2003).

Because the quantitative data showed few distinctions among teachers based on demographic characteristics, and virtually none for administrators, the interviewee selection process focused on having a range of school levels and genders represented by teachers and administrators (Table 4). Ultimately, the selection of interviewees came down to the combination of being recommended, the interviewee having interest in participating, and schedules aligning between the interviewee and the researcher.

Table 4

Interview Participants

Name	Position	School Level	Gender
Andrew	Administrator	High	M
Anna	Administrator	Elementary	F
Jake	Teacher	Middle	M
Kate	Teacher	High	F
Polly	Teacher	Middle	F
Shelley	Teacher	Elementary	F

Note. All teacher names are pseudonyms.

Instrumentation

Survey. The perception survey (Appendix A) for teachers, administrators, and policymakers was developed by generating possible questions for each group of stakeholders. After creating this list of questions, common themes were identified and questions were then grouped into six areas (Appendix B): perceptions of the current teacher evaluation systems, teacher evaluation process, organizational change, education reform, implementation of SB 191's teacher evaluation system, and SB 191's passage.

When designing a survey, it is imperative to develop "good" questions in order to receive consistent data collection (Fowler, 2009, p. 89). Good questions are defined as those with the following three properties: the questions are written clearly so that the responder is able to answer them, the questions mean the same thing to each respondent, and the questions have answer choices that are clear to each respondent (Fowler, 2009, pp. 89). To develop such questions, it is critical to avoid poorly defined terms and questions that ask more than one thing (Fowler, 2009). The standardization of responses

is essential (Fowler, 2009), as is providing primarily if not exclusively closed question responses (Fowler, 2009).

The questions for this survey primarily used a Likert scale format, which is considered an "appropriate" format for a survey focused on policies or perceptions (Fowler, 1995, p. 67). For the six sets of questions in the survey, there were four response options (Complete Disagree, Generally Disagree, Generally Agree, and Complete Agree) for four sets of questions and five response options for two sets of questions, with the extra option of "Don't Know" or "Not Applicable" being used sparingly as is recommended (University of Wisconsin-Madison, 2013). The fifth answer options were added to the two areas of questions where a participant may legitimately not know an answer or may find a question not applicable. For instance, a participant may not know if he/she will be actively participating in that school's development of the new teacher evaluation system, as a question in the fifth area asks. Similarly, a participant may find questions in the sixth area not applicable. For example, if a participant was not living in Colorado during the passage of SB 191, then he/she was unlikely to have been involved in the passage of the bill or have opinions about the media's coverage of the passage of the bill.

Having only four response options is considered a forced-choice rating scale, for there is no "Neither Disagree or Agree" response option and this can be seen as creating a bias by forcing individuals to have opinions (Friedman & Amoo, 1999); however, it is perceived as an acceptable format when individuals should have opinions on the matter (Friedman & Amoo, 1999). Phrasing that contains two dimensions in the response

columns, such as “Strongly Agree,” requires both an emotional dimension and a cognitive response (Fowler, 1995, p. 65); therefore, the answer options in the survey were solely in the cognitive dimension: “Completely Disagree, Generally Disagree, Generally Agree, and Completely Agree.” Additionally, the answer choices start with the negative options and move to positive because some studies have found that the reverse order (i.e., starting with “Completely Agree”) created false positives (Friedman & Amoo, 1999).

To assess the survey’s clarity of questions and ease of use, three pilot study sessions were conducted in three schools in three different school districts in the greater Denver metropolitan region. Pilot studies are critical to determining the value of the questions asked as well as the way in which they are asked (Fowler, 2009). A total of 17 individuals participated in the pilot studies of this survey. The first pilot study was a focus group that occurred at a school, consisted of six high school administrators, and lasted approximately 30 minutes. The first 10 minutes were devoted to individuals completing the draft survey and the final 20 minutes allocated to discussing participants’ perceptions of the survey. The second pilot study was also a focus group conducted in a school, but this group comprised five elementary school teachers. It lasted approximately 30 minutes as well, with the same time allocations as the first pilot focus group. The third and final pilot study had a total of six elementary school teachers and administrators, and was administered electronically at the participants’ convenience. The participants took the survey on his/her own time and later spoke with the researcher on the phone for approximately 10 minutes within a day of completing the survey. This second model was used because it both replicated the actual survey conditions and because it counteracted a

group discussion's "imperfect way to gather systemic information about the pretest experience" (Fowler, 2009, p. 123).

Members of all three pilot study groups were asked the same questions about their response to the survey (Appendix C). Participants from all three pilot studies provided recommendations for minor revisions and clarification of the survey. In the first pilot study, participants recommended the addition of a fifth column, which was added to questions in the fifth and sixth areas. Participants were asked about the "Generally Disagree" phrasing, and most found it acceptable with some even preferring it and none finding it off-putting. Based on feedback from individuals in the third and final pilot study, an open response area was added to the end of each section so that participants had the opportunity to share "additional information and/or comments" before clicking along to the next section. Following the conclusion of pilot testing, the survey was edited for flow, grammar, and overall consistency.

The aim for this study was to receive a response rate of 20% across all three stakeholders. Given the number of teachers, professional others, and administrators in RSD, combined with the 100 members of the 2013 Colorado General Assembly, the number of participants needed to reach an overall *n* of 20% was 759 (Teachers, which included the subcategory of "Other Licensed Professionals," $n = 719$; Administrators, $n = 20$; Policymakers, $n = 20$). This 20% response rate threshold was selected so that potential subgroups (i.e., elementary teachers, high school administrators, etc.) might be large enough to produce statistically meaningful results (Bobko, 2001).

Interviews. In explanatory sequential designs, the "initial quantitative phase connects into the data collection of the follow-up qualitative phase" (Creswell & Plano Clark, 2011, p. 221). Interviews are designed "to obtain a special kind of information" (Merriam, 2009, p. 88). The purpose of the interview in this study was to gather professional insights from selected teachers, administrators, and policymakers about the quantitative data results.

Interviews generally take one of three forms: highly structured, semi-structured, and unstructured (Merriam, 2009, p. 89). Semi-structured interviews can be defined as a "conversation" (Merriam, 2009, p. 88), but a conversation that is guided by particular objectives and issues to be explored. The semi-structured interview also provides freedom to explore participants' perspectives while still following a predetermined structure for the interview. The semi-structured interview was selected as the best qualitative data gathering process for this study for two reasons: 1) it would allow for the collection of interviewee's personal reflections on potentially sensitive subject matter, such as teacher evaluations (Briggs & Domingue, 2011) and analysis of peers' responses to the survey, in a safe and confidential environment (Creswell, 2007); 2) it would allow for in-depth explorations of data with the interviewees.

Prior to the start of all interviews, survey data were analyzed and used to develop questions that were responsive to the survey themes and findings. A total of 11 questions were developed (Appendix E), with the final question being an open-response opportunity for the interviewee. The first 10 questions were based on specific data culled from survey analysis.

In following CBR's social action approach to research (Strand et al., 2003; Couto, 2003), the interviews were designed to collect insights about how current practitioners understood the data results as well as how they anticipated these data might be used constructively in the district to facilitate the implementation of the new teacher evaluations.

Data Collection

Survey. The survey was disseminated through Qualtrics, a University of Denver software program, able to handle large quantities of data and transfer the data into Statistical Package for the Social Sciences (SPSS) software for data analysis. The survey was distributed online through professional email accounts; there were no paper copies of the survey. The benefits of using an electronic survey included low costs, quick returns and relatively easy data computations, which made this the best method of disseminating the survey for this study despite the potential for low response rates (Nulty, 2008).

RSD disseminated the surveys to teachers and administrators and sent follow-up emails encouraging participation. Dissemination from the district ensured all recipients that the survey had district approval. The survey window was 14 days. Two reminder emails were sent during this window. For policymakers, a member of the Senate Education Committee sent out the survey to members of Colorado's Senate and House of Representative Education Committees. These policymakers were asked to complete the survey and send the survey link to colleagues in other committees, thereby utilizing the snowball sampling method (Merriam, 2009). The window for the policymakers' survey was originally eight days due to the then-impending start of the legislative session, but

later expanded to the same 14 days that RSD teachers and administrators received. A total of two reminder emails were sent during the 14-day window.

Interviews. Interviews were conducted in a quiet space (Creswell, 2007) on the phone over the course of three weeks. Each interview was recorded and transcribed, with the interviewee's consent. Interviews were limited to 30 minutes.

In conjunction with an RSD assistant superintendent's recommendation of exemplary teachers and administrators and a current state government official's recommendation, eight individuals were emailed and asked to participate in an interview. A total of six interviews were conducted: four with RSD teachers and two with RSD administrators. It was not possible to interview a policymaker.

Confidentiality

The survey was anonymous and the interviews were confidential. This was made explicit to participants in the consent forms (Appendixes D, F, and G). Survey participants were not asked to provide an identifiable personal name or data in the demographics section of the survey. For interviewees, it was of utmost importance that they felt they were in safe environment, which was made possible by the selection of non-public settings for interviews. Such individuals were also protected in the written form of this research through the use of code names. The researcher was the only individual with access to the code key. No individual's real name or school name appears in the study.

Data Analysis

The data analysis took place in two distinct stages.

Survey. Survey results were analyzed first. The data were downloaded electronically from Qualtrics to Excel to SPSS. The researcher ran descriptive statistics, primarily frequencies, Pearson chi-square tests, and one-way ANOVAs with the Scheffé and Tukey post-hoc tests in SPSS. Frequency tests were used both to determine survey participant populations based on demographic data and to gain insights into participants' perspectives on the few non-Likert questions asked. The Pearson chi-square tests were used to determine how members of specific groups responded to individual questions, and then to assess if there were statistically significantly different responses by group. Because the Likert scale used in this study had only four or five answer choices, the data were not continuous. With non-continuous data, certain common analyses, such as *t*-tests, were inappropriate. When the Pearson chi-square test did reveal statistically significantly different item results between three or more groups, these questions were later run through one-way ANOVAs with Scheffé and Tukey post-hoc tests used to determine precisely which group(s) differed from the others. For all tests in this study, the alpha level was set at .05, the alpha level used most frequently in social science research (Web Center for Social Research Methods, 2013).

With each grouping, there were 38 Likert questions run through the Pearson chi-square tests plus an additional two (PR, Primary responsibility of implementation, and CF, Central Focus of SB 191), for a total of 40 questions. Ultimately, eight different sets of groupings were run through the Pearson chi-square tests for the 40 questions (see Appendix H for a condensed list). The sets were: Teacher, Administrator, Policymaker;

Teachers by School Level, Gender, Educator Type, License Type, and Ethnicity; and Administrators by School Level and Gender.

Interviews. The interview data analysis took place upon the completion of qualitative data collection. For explanatory sequential designs, this second analysis phase addresses "whether and how the qualitative data help to explain the quantitative results" (Creswell & Plano Clark, 2011, p. 221). The interviews were transcribed by the researcher during the first round of analysis (Merriam, 2009). The transcribed interviews were then reviewed in the program ATLAS.ti using the constant comparative method, which enabled the researcher "to identify patterns" in results (Merriam, 2009, pp. 30-31). The objective was to look for themes, poignant quotes, and concepts that helped explain the quantitative data.

Limitations

This was a study of teachers and administrators within a single Colorado school district, in collaboration with state-level policymakers. The results from teachers and administrators may not be transferred to educators across the state. Other limitations to this study fall into four categories: explanatory methodology, survey limitations in general, additional limitations for policymakers, and survey response rates.

One limitation of this study was the design methodology: By choosing an explanatory sequential design, the researcher began with a survey of her design and ended with interviews. Therefore, the questions were developed by the researcher instead of originating with the key stakeholder groups themselves. This is not ideal from the

CBR perspective (Reason & Bradbury, 2001), for the questions are the researcher's own instead of coming from the collective inquiry of both researcher and practitioners.

Another limitation of this study was the decision to disseminate the survey through the internet and professional email accounts only. Not all participants may have had private electronic access to the survey and this may have influenced the survey results or the return rate of the survey. Another concern was that, for some participants, the email containing the survey could have gone into the spam filter for their email inbox. Finally, although online surveys are significantly easier to administer, they generally have a weak response rate compared to paper-surveys (Nulty, 2003). This survey was no exception given that the response rate averaged across all three groups was 18%, short of the desired 20%.

Additionally, the subject matter was personal, and this may have discouraged individuals from taking the survey (Briggs & Domingue, 2011). Or, in related fashion, the personal nature of the subject matter may have created the potential for bias in results, particularly for the open-ended questions at the conclusion of each set of questions. While the responses gave insights into survey participants' perspectives, the response rates of survey participants in general, let alone those who choose to add personal comments, made it more likely that those who wrote had particularly strong feelings about matters related to SB 191, in one direction or another, rather than being reflective of all RSD teachers and administrators or all statewide legislators.

In addition to these general limitations of online survey dissemination, the survey to policymakers was limited by its dissemination method and timing. The use of snowball

sampling for disseminating the survey to members of the General Assembly was ineffective, as illustrated by the low response rate of 12%. Although snowball sampling is a legitimate way to gather data (Merriam, 2009), it was unsuccessful in this study because few policymakers took the survey and it is unclear if they sent the survey to colleagues in other committees. Similarly, survey dissemination was limited because only members of Senate and House of Representatives Education Committees received a formal introduction to the survey, whereas all teachers and administrators in RSD received official introductory email letters. Finally, the timing of the survey dissemination to members of the General Assembly was yet another limitation because it occurred within days of the start of the 2013 legislative session. Ultimately, the response rate for policymakers was so low as to require only limited analysis of policymaker data, for there was insufficient data from which to determine trends in perceptions. Despite the low response rate, the researcher had intended to interview one member of the legislative assembly, but was unable to do so due to a lack of timely communication with a policymaker.

The survey response rates were a limitation for three reasons. First, the response rates, based on the total number of survey participants who selected "Teacher," "Administrator" or "Policymaker" for the first question, were low (see Table 2) for teachers and policymakers. Second, examination of individual questions revealed even lower participation rates, given that the greatest number of teachers to answer any single question was 562 and the greatest number of policymakers to answer any single question was 9. This would bring the "true" response rate of teachers to 16% and of policymakers

to 9%. Similarly, the calculation rate of RSD participant response rates by gender and school level (Table 2) had to be done without the data of 102 teachers and 11 administrators who failed to note their gender and/or school level. Thus, it was challenging to create a completely accurate list of survey participants. Third and finally, the response rates for individuals in the "Mental Health" and "Nurses" category were exceptionally low; however, it is still uncertain whether individuals in these categories will be evaluated using the new teacher evaluation system, so these low rates should not have adversely skewed the data. However, they did lower the percentage rates for responses.

Researcher Bias

In line with the constructivist perspective, there is inherent researcher bias in this study. As a former high school teacher and administrator, as well as the daughter and sister of teachers and administrators, the researcher has an appreciation of and affinity for these professionals. She believes that ground-level educators often have meaningful insights into the reality of reforming schools, an idea that was a driving force in this study. The researcher also had relationships with members of the RSD community prior to the start of this study because she previously worked in one of the RSD schools. Additionally, the researcher had worked with many of the state legislators who sponsored SB 191. She also had volunteered for one of the state legislators who sponsored the bill, though she was not involved with the writing and passage of the bill in any way.

Despite the researcher's personal involvement in education and Colorado state politics, the researcher strove to keep all personal biases out of the research collection

and analysis. She did this by generating topic areas to explore in the survey and interviews, pre-testing the survey with three separate pilot groups to ensure that the questions were clear, and using the same survey with all three key stakeholder groups. To control for bias, the researcher followed professional processes for data collection and analysis.

Summary

In using an explanatory sequential design approach, this study delved into how teachers, administrators, and policymakers perceived their role in the implementation of SB 191's new teacher evaluations through both quantitative and qualitative measures. The quantitative approach consisted of the administration of a survey instrument to a large sample of teachers, administrators, and policymakers. The qualitative approach was composed of follow up interviews with individuals who were asked to expand on patterns in, and add detail to, the quantitative results. The mixed methods design enabled this study to have both breadth and depth in understanding how key education stakeholders perceived their role in the implementation of SB 191.

Chapter Four: Findings

Introduction

This section contains the findings from both the quantitative and qualitative data collected to address this study's research questions. The research questions were:

1) How do Colorado's teachers, administrators, and policymakers perceive their respective roles in the implementation of SB 191's new teacher evaluations?

2) Do Colorado's teachers', administrators', and policymakers' responses vary by demographic variables?

3) How do the responses of Colorado's teachers, administrators, and policymakers compare?

Quantitative Findings

The quantitative findings come from survey results. Because the objective of the survey data was to understand how the three stakeholder groups each perceived its role in the implementation of SB 191, as well as perceptions within subgroups of teachers and administrators and a comparison of the three stakeholder groups, the analysis focused on comparing how members of each group responded to the questions. With the Likert-like questions used in the survey, there were only four (or occasionally five) answer choices instead of continuous answers. The best means to analyze responses to these non-continuous variables was to use the Pearson chi-square test for 40 questions (38 Likert questions plus two others), followed with one-way ANOVAs with the Scheffé post-hoc

tests when statistically significant differences occurred among three or more groups on a single question (see Appendix H for a condensed list of questions and labels). For all questions, p = .05. The data provided are percentages of collective responses on the agreed response spectrum (Generally Agree plus Completely Agree).

Survey analysis by group.

Teachers. The survey questions were run six times for teachers: first when teachers' responses were compared with administrators' and policymakers' responses, and five times with different teacher subgroups. The subgroups of teachers for additional analysis were: school level, gender, type of teacher (classroom teacher of a TCAP-tested subject, classroom teacher of a non TCAP-tested subject, and Other Licensed Professional), teacher license type (Non-Professional, which was Initial or Alternative, and Professional), and ethnicity (minority and white). Exploration of teachers' responses to the survey questions in these five subgroupings through Pearson chi-square tests allowed for a richer understanding of the data.

The comparison of teachers' collective responses to the 38 Likert questions (Tables 5-10) and two other questions (Tables 11-14) by the five teacher subgroupings illustrated that, overall, teachers had similar responses to all questions regarding teacher evaluations, education reform, SB 191, and change. However, closer inspection revealed that teachers in the non-professional and minority subgroups most frequently differed from their peers by often expressing greater agreement with various questions, particularly in the last three areas of questions. The small number of participants in these groups, approximately 40 in both, may have contributed to how their results could

Table 5

Current Teacher Evaluation Questions, by Teachers and all Teacher Subgroups

Label	Question		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP - Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
CTE_1	Current TE improved your teaching.	69%	71%	66%	65%	65%	64%	69%	72%	68%	63%	85%	68%	67%	68%
CTE_2	Current TE has provided me with meaningful feedback.	71%	73%	68%	68%	65%	72%	70%	73%	72%	64%	78%	70%	66%	70%
CTE_3	Current TE has made me a better teacher overall.	65%	68%	61%	63%	65%	63%	65%	68%	65%	59%	78%	64%	70%	64%
CTE_1a	Current TE encourages professional growth for teachers.	70%	74%	60%	68%	70%	63%	70%	70%	69%	70%	73%	69%	71%	69%
CTE_2a	Current TE is linked to student learning.	70%	76%	64%	65%	65%	65%	70%	70%	70%	72%	75%	69%	76%	69%
CTE_3a	Current TE is able to assess teachers' overall teaching abilities.	62%	66%	57%	57%	65%	57%	63%	64%	60%	62%	68%	61%	64%	61%
CTE_4a	Current TE is a high stress process for teachers.	64%	69%	61%	58%	75%	60%	65%	63%	64%	65%	58%	65%	69%	64%
CTE_5a	Current TE is a high stress process for administrators.	53%	62%	49%	39%	55%	50%	54%	52%	49%	57%	33%	54%	65%	52%
CTE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	44%	50%	40%	35%	45%	35%	46%	42%	44%	49%	58%	42%	49%	43%
CTE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	55%	59%	53%	50%	53%	48%	56%	54%	54%	59%	65%	54%	62%	54%
	Counts	547	235	114	144	20	107	403	265	153	132	40	471	45	465

Note: All counts an average.

Table 6

Ideal Teacher Evaluation Questions, by Teachers and all Teacher Subgroups

Label	Question: Ideal teacher evaluations:		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP-Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
IdealTE_1	distinguish strong teachers from weak teachers.	59%	57%	63%	56%	79%	56%	60%	62%	48%	66%	53%	59%	37%	61%
IdealTE_2	help teachers become better teachers.	94%	93%	95%	95%	90%	93%	94%	95%	92%	94%	95%	94%	93%	94%
IdealTE_3	have the ability to improve student learning.	93%	92%	95%	92%	95%	89%	93%	94%	90%	91%	90%	93%	91%	92%
IdealTE_4	are linked to student assessment data of any form.	66%	67%	68%	62%	60%	55%	69%	70%	61%	65%	77%	65%	67%	66%
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	31%	30%	39%	24%	30%	27%	31%	36%	23%	32%	35%	30%	22%	31%
IdealTE_6	are linked to student assessment data created by the teacher.	79%	78%	82%	76%	85%	74%	80%	80%	76%	82%	80%	78%	82%	79%
	Counts	542	237	114	144	20	108	406	263	151	127	40	472	45	467

Note: All counts an average.

Table 7

Education Reform Questions, by Teachers and all Teacher Subgroups

Label	Question		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP - Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
EdRef_1	EdRef: Education reforms improve student learning.	55%	62%	57%	38%	70%	47%	57%	54%	51%	62%	73%	53%	62%	54%
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	34%	41%	32%	23%	45%	34%	34%	34%	36%	35%	55%	32%	32%	34%
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	55%	54%	59%	52%	65%	56%	55%	58%	53%	52%	70%	54%	64%	54%
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	4%	6%	1%	3%	0%	3%	4%	4%	3%	7%	8%	3%	11%	3%
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	7%	7%	5%	6%	5%	5%	7%	5%	9%	8%	15%	6%	9%	6%
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	68%	76%	62%	60%	65%	63%	70%	72%	63%	68%	70%	68%	69%	68%
	Counts	531	236	114	144	20	108	405	256	152	122	40	471	45	466

Note: All counts an average.

Table 8

SB 191 Teacher Evaluation Questions, by Teachers and all Teacher Subgroups

Label	Question: Ideal teacher evaluations:		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP - Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	46%	50%	39%	45%	45%	47%	46%	46%	51%	41%	43%	46%	49%	45%
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	13%	14%	16%	12%	5%	14%	13%	13%	11%	16%	38%	11%	24%	12%
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	15%	13%	18%	17%	10%	19%	14%	15%	14%	18%	35%	13%	22%	14%
	Counts	528	236	114	145	20	108	405	255	150	122	40	472	45	466

Note: All counts an average.

Table 9

SB 191 General Questions, by Teachers and all Teacher Subgroups

Label	Question: Ideal teacher evaluations:		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP - Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
SB191_1	I have a strong understanding of the legislation SB 191.	47%	43%	44%	55%	45%	59%	44%	46%	54%	40%	38%	47%	49%	46%
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	3%	3%	4%	2%	5%	4%	3%	3%	4%	3%	3%	3%	4%	3%
SB191_3	I supported SB 191 during its development.	14%	15%	18%	10%	5%	16%	14%	15%	15%	11%	18%	14%	18%	14%
SB191_4	I was pleased with the final version of SB 191 as it was passed.	14%	15%	16%	11%	10%	19%	13%	13%	16%	14%	28%	13%	16%	14%
SB191_5	The media portrayed teachers fairly during the bill's passage.	9%	10%	11%	6%	5%	9%	9%	9%	10%	8%	33%	7%	13%	8%
SB191_6	The media portrayed administrators fairly during the bill's passage.	16%	14%	19%	15%	20%	18%	15%	16%	18%	13%	33%	14%	31%	14%
SB191_7	The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.	32%	35%	34%	28%	25%	36%	31%	36%	27%	30%	43%	32%	40%	31%
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	36%	38%	41%	30%	35%	34%	37%	38%	34%	36%	45%	36%	31%	37%
	Counts	524	235	113	145	20	108	402	252	145	120	40	469	45	464

Note: All counts an average.

Table 10

Change Questions, by Teachers and all Teacher Subgroups

Label	Question: Ideal teacher evaluations:		School Level				Gender		Teacher Type			License Type		Ethnicity	
		Teachers	Elem.	Middle	High	K-8 + Other	M	F	TCAP - Teacher	non-TCAP Teacher	Other Licensed Pro.	Non-Pro.	Pro.	M	W
Change_1	I view change as an opportunity for growth.	96%	97%	97%	92%	95%	94%	96%	95%	95%	98%	98%	95%	100%	95%
Change_2	I respond well when policies change at work.	77%	78%	80%	72%	85%	69%	80%	78%	72%	81%	70%	78%	78%	78%
Change_3	I respond well to change at work when I am involved with creating the change.	98%	98%	98%	97%	100%	97%	98%	98%	97%	98%	100%	98%	98%	98%
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	82%	83%	86%	77%	75%	75%	84%	83%	82%	79%	80%	81%	89%	81%
Change_5	There is a right way and a wrong way to introduce new policies in a school.	94%	92%	97%	97%	90%	98%	93%	95%	92%	93%	93%	95%	93%	94%
	Counts	538	237	113	145	20	108	405	260	152	126	40	472	45	467

Note: All counts an average.

Table 11

Perceived Central Focus of SB 191 Reforms (CF), By Teachers and Teacher Subgroups

Teachers and Teacher Subgroups	Students	Teachers	Admin	Policymakers	Total #s
All Teachers	16%	59%	1%	24%	504
Elementary	17%	57%	1%	25%	228
Middle	16%	63%	1%	20%	110
High	13%	59%	1%	27%	143
K-8 + Other	11%	72%	6%	11%	18
					499
Males	9%	63%	2%	26%	108
Females	18%	58%	1%	23%	388
					496
TCAP	14%	63%	1%	22%	243
non-TCAP	19%	49%	3%	29%	144
Other Licensed Professional	16%	60%	1%	23%	116
					503
Non-Professional License	26%	64%	3%	8%	39
Professional License	15%	59%	1%	25%	456
					495
Minority	28%	49%	0%	23%	43
White	15%	60%	2%	24%	452
					495

Table 12

Pearson Chi-Square data for CF (see Table 11), by Teacher Subgroups

Teacher Subgroup	Value	df	Asymp. Sig (2-sided)
School Level	14.214	15	.509
Gender	10.650	5	.059
Education Type	15.059	10	.130
License Type	3.314	5	.652
Ethnicity	4.349	5	.500

Note: For all questions, at least three cells had an expected count less than five, which may have impacted significance.

Table 13

Which group has primary responsibility for successful implementation of new teacher evaluations?, By Teachers and Teacher Subgroups

Teachers and Teacher Subgroups	Teachers	Admin	District Personnel	Policy makers	Other	Unsure	Total #s
All Teachers	13%	56%	14%	4%	1%	13%	528
Elementary	9%	61%	15%	2%	1%	12%	237
Middle	18%	54%	14%	5%	0%	9%	114
High	14%	51%	14%	5%	1%	16%	144
K-8 + Other	15%	60%	10%	5%	0%	10%	20
							515
Males	16%	47%	12%	8%	1%	17%	107
Females	12%	59%	15%	3%	1%	11%	406
							513
TCAP	12%	59%	11%	4%	0%	13%	255
non-TCAP	16%	53%	15%	5%	1%	9%	150
Other Licensed Professional	9%	54%	19%	2%	0%	16%	122
							527
Non-Professional License	20%	55%	13%	5%	0%	8%	40
Professional License	12%	57%	14%	4%	1%	13%	472
							512
Minority	9%	56%	11%	7%	2%	16%	45
White	13%	56%	15%	4%	0%	12%	467
							512

Table 14

Pearson Chi-Square data for PR (see Table 13), by Teacher Subgroups

Teacher Subgroup	Value	df	Asymp. Sig (2-sided)
School Level	6.760	9	.662
Gender	4.723	3	.193
Education Type	9.323	6	.156
License Type	7.838	3	.049
Ethnicity	5.882	3	.117

Note: For all questions, at least one cell had an expected count less than five, which may have impacted significance.

regularly differ from the overall perceptions of teachers who took the survey. The three other groups whose perceptions occasionally differed from the majority perception by 8% or more were male teachers and teachers in the High School and K-8 + Other schooling groups.

When asked to reflect on which group was the primary focus of SB 191, the overwhelming conclusion was teachers, although teachers with a non-Professional License and of a minority ethnicity believed the reform was designed more for students (Table 11). When asked who had the greatest responsibility in implementation, the answer was administrators (Table 13).

The results of the Pearson chi-square tests for each of the teacher subgroups (Appendixes J-N), and accompanying one-way ANOVA with Scheffé post-hoc tests for the School Level and Educator Type subgroups (Appendixes O and P, respectively), collectively illustrated that the subgroupings with the greatest number of statistically significant differences among groups members were school level and license type (Table 15). The ANOVA results illustrated that, within the school level grouping, high school

Table 15
Teacher Subgroup Analyses

	# of Questions, total	# of Questions with Statistically Significantly Different Results	# of Questions with Statistically Significantly Different Results in ANOVA	% of Questions with Stat Sig Dif
School Level	40	15	12	30.0%
Gender	40	5	–	12.5%
Educator Type	40	3	2	5.0%
License Type	40	11	–	27.5%
Ethnicity	40	5	–	12.5%

teachers most frequently differed from their peers. For certain questions, such as one in the Educator Type subgroup, the running of a one-way ANOVA with post-hoc Scheffé test resulted in non-statistically significant findings among the groups, indicating that the original statistically significant finding was most likely due to correlations within the subgroupings.

In total, the majority of teachers responded uniformly on issues related to teacher evaluations, education reforms, SB 191, and change. The data indicated that RSD teachers were largely comfortable with the current teacher evaluation system (C-TE_1a) and its connection to student learning (C-TE_2a), but skeptical of education reforms improving student learning in general (EdRef_1) and doubtful that SB 191's new teacher evaluations will improve teachers' practices (SB191TE_2) or student learning (SB191TE_3) and achieve its desired impact (SB191_8). Additionally, teachers appeared uncertain about the idea of teacher evaluations separating strong teachers from weak teachers (IdealTE_1), decidedly uncomfortable with the idea of linking state-created student assessments to teacher evaluations (IdealTE_5), yet comfortable linking teacher-made assessments to teacher evaluations (IdealTE_6). The consistent outliers to these general trends were teachers with non-professional licenses and of a minority background.

Administrators. The survey data for administrators were run three times: first against teachers' and policymakers' responses, second internally by school level (Appendix Q), and third internally by gender (Appendix R). Analyzing the data from administrators in these additional subgroupings allowed for greater insight into

administrators' perspectives on the survey questions. The comparison of administrators' collective responses to the 38 Likert questions (Tables 16-21) and two other questions (Tables 22-25) by the two administrator subgroupings illustrated that administrators

Table 16

Current Teacher Evaluations, by Administrators and Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
CTE_1	Current TE improved your teaching.	–	–	–	–	–	–	–
CTE_2	Current TE has provided me with meaningful feedback.	–	–	–	–	–	–	–
CTE_3	Current TE has made me a better teacher overall.	–	–	–	–	–	–	–
CTE_1a	Current TE encourages professional growth for teachers.	90%	95%	79%	100%	67%	90%	91%
CTE_2a	Current TE is linked to student learning.	80%	81%	71%	86%	67%	74%	82%
CTE_3a	Current TE is able to assess teachers' overall teaching abilities.	86%	91%	79%	100%	67%	90%	88%
CTE_4a	Current TE is a high stress process for teachers.	52%	38%	57%	69%	33%	61%	44%
CTE_5a	Current TE is a high stress process for administrators.	34%	29%	29%	43%	33%	42%	26%
CTE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	68%	71%	64%	79%	67%	63%	76%
CTE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	71%	67%	71%	86%	67%	68%	77%
	Counts	59	21	14	14	3	19	34

Note: All counts an average.

Table 17

Ideal Teacher Evaluations, by Administrators and all Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
IdealTE_1	distinguish strong teachers from weak teachers.	76%	86%	71%	79%	100%	74%	82%
IdealTE_2	help teachers become better teachers.	97%	95%	93%	100%	100%	100%	94%
IdealTE_3	have the ability to improve student learning.	97%	95%	100%	93%	100%	95%	97%
IdealTE_4	are linked to student assessment data of any form.	86%	81%	93%	79%	100%	95%	79%
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	72%	76%	57%	72%	100%	79%	68%
IdealTE_6	are linked to student assessment data created by the teacher.	79%	76%	86%	69%	100%	90%	73%
	Counts	58	21	14	14	3	19	34

Note: All counts an average.

Table 18

Education Reform, by Administrators and all Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
EdRef_1	EdRef: Education reforms improve student learning.	75%	86%	50%	86%	67%	68%	79%
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	57%	62%	36%	64%	67%	53%	59%
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	66%	67%	50%	71%	67%	63%	65%
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	7%	10%	0%	14%	0%	11%	6%
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	7%	10%	0%	14%	0%	11%	6%
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	66%	71%	57%	57%	100%	58%	71%
	Counts	56	21	14	14	3	19	34

Note: All counts an average.

Table 19

SB 191 Teacher Evaluations, by Administrators and all Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	89%	95%	86%	86%	67%	100%	82%
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	39%	57%	29%	29%	33%	42%	38%
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	32%	48%	21%	21%	33%	26%	35%
	Counts	54	21	14	14	3	19	34

Note: All counts an average.

Table 20

SB 191, by Administrators and all Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
SB191_1	I have a strong understanding of the legislation SB 191.	83%	95%	71%	79%	67%	74%	88%
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	17%	19%	7%	21%	33%	26%	12%
SB191_3	I supported SB 191 during its development.	50%	62%	43%	43%	33%	63%	44%
SB191_4	I was pleased with the final version of SB 191 as it was passed.	56%	71%	29%	57%	67%	58%	56%
SB191_5	The media portrayed teachers fairly during the bill's passage.	23%	29%	14%	23%	0%	22%	21%
SB191_6	The media portrayed administrators fairly during the bill's passage.	28%	43%	14%	23%	0%	22%	29%
SB191_7	The media portrayed policymakers (i.e., legislators, lobbyists) fairly during the bill's passage.	40%	48%	46%	39%	0%	33%	46%
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	83%	95%	71%	79%	67%	84%	82%
	Counts	54	21	14	14	3	19	34

Note: All counts an average.

Table 21

Change, by Administrators and all Administrator Subgroups

Label	Question		School Level				Gender	
		Admin	Elem.	Middle	High	K-8 + Other	M	F
Change_1	I view change as an opportunity for growth.	100%	100%	100%	100%	100%	100%	100%
Change_2	I respond well when policies change at work.	98%	100%	93%	100%	100%	100%	97%
Change_3	I respond well to change at work when I am involved with creating the change.	100%	100%	100%	100%	100%	100%	100%
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	97%	95%	100%	93%	100%	100%	94%
Change_5	There is a right way and a wrong way to introduce new policies in a school.	93%	91%	100%	93%	100%	95%	94%
	Counts	58	21	14	14	3	19	34

Note: All counts an average.

Table 22

Perceived Central Focus of SB 191 Reforms (CF), By Administrators and Administrator Subgroups

Administrators and Administrators Subgroups	Students	Teachers	Admin	Policymakers	Total #s
All Administrators	43%	45%	2%	9%	53
Elementary	48%	48%	0%	5%	21
Middle	29%	50%	7%	14%	14
High	54%	39%	0%	8%	13
K-8 + Other	33%	33%	0%	33%	3
					51
Males	37%	58%	0%	5%	19
Females	46%	39%	3%	12%	33
					52

Table 23

Pearson Chi-Square data for CF (see Table 22), by Administrator Subgroups

Administrator Subgroup	Value	df	Asymp. Sig (2-sided)
School Level	6.749	9	.663
Gender	2.271	3	.518

Note: For all questions, at least four cells had an expected count less than five, which may have impacted significance.

Table 24

Which group has primary responsibility for successful implementation of new teacher evaluations?, By Administrators and Administrator Subgroups

Administrators and Administrators Subgroups	Teachers	Admin	District Personnel	Policy makers	Other	Unsure	Total #s
All Administrators	4%	75%	18%	0%	0%	4%	55
Elementary	5%	86%	10%	0%	0%	0%	21
Middle	0%	79%	21%	0%	0%	0%	14
High	0%	64%	21%	0%	0%	14%	14
K-8 + Other	33%	33%	33%	0%	0%	0%	3
							52
Males	0%	84%	16%	0%	0%	0%	19
Females	6%	68%	21%	0%	0%	6%	34
							53

Table 25

Pearson Chi-Square data for PR (see Table 24), by Administrator Subgroups

Administrator Subgroup	Value	df	Asymp. Sig (2-sided)
School Level	16.042	9	.066
Gender	2.838	3	.417

Note: For all questions, at least five cells had an expected count less than five, which may have impacted significance.

responded in a similar fashion to all questions reviewed. The one subgroup that frequently had varying responses was the K-8+Other group; however, this subgroup

consisted of three individuals and so these results do not indicate significant differences among administrators by subgroup.

An analysis of administrators' collective responses compared with the administrators' subgrouping responses showed that RSD administrators had variations in their responses but few statistically significant differences within subgroups (Table 26).

Table 26

Administrator Subgroup Analyses

	# of Questions, total	# of Questions with Statistically Significantly Different Results	# of Questions with Statistically Significantly Different Results in ANOVA	% of Questions with Stat Sig Dif
School Level	37	6	1	2.7%
Gender	37	1	–	2.7%

For those few questions that required ANOVA analysis (Appendix S), the only question that resulted in a statistically significant difference was between Elementary and Middle school administrators (SB191_8). The implication of having only two statistically significant differences between administrator subgroups was that, at least when divided by school level and gender, RSD administrators' had similar perceptions on the issues presented in the survey.

Collectively, administrators expressed faith that the current teacher evaluations have supported teachers (C-TE_1a, C-TE_3a) and were connected to student learning (C-TE_2a). Administrators believed education reforms are designed to improve student learning (EdRef_1) and that SB 191 will not be an exception (SB191_8). Though not many administrators were involved with the development of SB 191 (SB191_2), the majority had a strong understanding of it (SB191_1) and about half were pleased with it

(SB191_4). Administrators believed that they will play an active role in the implementation of SB 191 in their schools (SB191TE_1), but exhibited considerable uncertainty as to how teachers in their building feel about SB 191 (SB191TE_2, SB191TE_3).

Policymakers. The survey data for policymakers were run one time alongside teachers' and administrators' collective perceptions (see Tables 28-35 in the following section). The low response rate of policymakers prevented any additional survey analysis. Additionally, the low level of response necessitates that one interpret the results with caution, for the 9% who responded represent, in fact, the perceptions of nine people and are unlikely to be representative of the collective perceptions of Colorado's General Assembly.

In spite of the low response rate from policymakers, the data were illustrative of how some members of the General Assembly perceived the implementation of SB 191 and can be used to start exploring how policymakers' perceptions compare to those of teachers and administrators. The data results for policymakers were best seen in comparison to teachers and administrators in the following section.

Comparing responses of teachers, administrators, and policymakers. Teachers', administrators', and policymakers' responses, and subgroupings of responses as was appropriate, illustrated collective perceptions of individuals within all three groups. The juxtaposition of teachers', administrators', and policymakers' (TAP) responses to the survey questions in the six areas of questioning through the Pearson chi-

square analyses and subsequent one-way ANOVA added another layer of analysis as it illuminated differences in perceptions among the three groups.

Unlike the analyses of subgroups of teachers and subgroups of administrators, the comparison of teachers', administrators' and policymakers' responses to the 40 questions revealed a number of statistically significantly differences in perception among these three groups (Table 27). As Tables 28-35 show, the differences in perceptions among

Table 27
Review of Questions with Needed Analyses by Teachers, Administrators, and Policymakers (TAP)

	# of Questions, total	# of Questions with Statistically Significantly Different Results	# of Questions with Statistically Significantly Different Results in ANOVA	% of Questions with Stat Sig Dif
Teachers, Administrators, and Policymakers	27	19	15	56%
Teachers and Administrators Only	10	7	–	70%
Teachers Only	3	–	–	–

members of these three groups permeated all areas of questioning. Additionally, the one-way ANOVA with Scheffé post-hoc test results for the 19 questions with statistically significant results (Appendix T) among the three key stakeholders revealed that teachers and administrators had statistically significant differences 15 times as well as teachers and policymakers for one question (SB191_2) and administrators and policymakers for one question (Change_4). The survey results indicated that teachers' and administrators'

Table 28

Current Teacher Evaluation System, by TAP

Label	Question	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	69%	—	—	—	—	—
C-TE_2	Current TE has provided me with meaningful feedback.	71%	—	—	—	—	—
C-TE_3	Current TE has made me a better teacher overall.	65%	—	—	—	—	—
C-TE_1a	Current TE encourages professional growth for teachers.	70%	90%	—	14.089	3	.003
C-TE_2a	Current TE is linked to student learning.	70%	80%	—	3.975	3	.264
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	62%	86%	—	13.935	3	.003
C-TE_4a	Current TE is a high stress process for teachers.	64%	52%	—	4.894	3	.180
C-TE_5a	Current TE is a high stress process for administrators.	53%	34%	—	7.929	3	.048
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	44%	68%	—	14.267	3	.003
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	55%	71%	—	7.616	3	.055
	Counts	547	59	—	—	—	—

Note: Counts are an average for Teachers and Administrators. For questions C-TE_1a through C-TE_7a, at least one cell had an expected count less than five, which may have impacted significance.

Table 29

Ideal Teacher Evaluations, by TAP

Label	Question: Ideal teacher evaluations:	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	59%	76%	78%	16.071	6	.013
IdealTE_2	help teachers become better teachers.	94%	97%	100%	6.057	6	.417
IdealTE_3	have the ability to improve student learning.	93%	97%	100%	9.363	6	.154
IdealTE_4	are linked to student assessment data of any form.	66%	86%	78%	31.742	6	.000
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	31%	72%	56%	47.296	6	.000
IdealTE_6	are linked to student assessment data created by the teacher.	79%	79%	56%	6.728	6	.347
	Counts	542	58	9	—	—	—

Note. Count is average for Teachers. For all questions, at least four cells had an expected count less than five, which may have impacted significance.

Table 30

Education Reform, by TAP

Label	Question	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	55%	75%	56%	44.739	6	.000
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	34%	57%	56%	42.764	6	.000
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	55%	66%	67%	5.113	6	.529
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	4%	7%	22%	41.184	6	.000
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	7%	7%	33%	32.347	6	.000
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	68%	66%	33%	8.476	6	.205
	Counts	531	56	9	–	–	–

Note: Counts are an average for Teachers and Administrators. For questions EdRef_1 through EdRef_6, at least five cell had an expected count less than five, which may have impacted significance.

Table 31

SB191 Teacher Evaluations, by TAP

Label	Question	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	46%	89%	—	119.076	4	.000
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	13%	39%	—	47.820	4	.000
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	15%	32%	—	37.180	4	.000
	Counts	528	54	—	—	—	—

Note: Counts are an average for Teachers and Administrators. For all questions, at least one cell had an expected count less than five, which may have impacted significance.

Table 32

SB 191, by TAP

Label	Question	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	47%	83%	67%	49.723	8	.000
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	3%	17%	33%	45.832	8	.000
SB191_3	I supported SB 191 during its development.	14%	50%	56%	85.960	8	.000
SB191_4	I was pleased with the final version of SB 191 as it was passed.	14%	56%	56%	77.199	8	.000
SB191_5	The media portrayed teachers fairly during the bill's passage.	9%	23%	22%	24.294	8	.002
SB191_6	The media portrayed administrators fairly during the bill's passage.	16%	28%	44%	21.443	8	.006
SB191_7	The media portrayed policymakers (i.e., legislators, lobbyists) fairly during the bill's passage.	32%	40%	44%	14.431	8	.071
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	36%	83%	67%	62.702	8	.000
	Counts	524	54	9	–	–	–

Note: Counts are an average for Teachers and Administrators. For all questions, at least six cells had an expected count less than five, which may have impacted significance.

Table 33

Change, by TAP

Label	Question	Teachers	Admin	Policy makers	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	96%	100%	100%	17.552	6	.007
Change_2	I respond well when policies change at work.	77%	98%	78%	29.555	6	.000
Change_3	I respond well to change at work when I am involved with creating the change.	98%	100%	100%	9.351	6	.053
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	82%	97%	67%	25.191	6	.000
Change_5	There is a right way and a wrong way to introduce new policies in a school.	94%	93%	100%	4.232	6	.645
	Counts	538	58	9	–	–	–

Note: Counts are an average for Teachers and Administrators. For all questions, at least four cells had an expected count less than five, which may have impacted significance.

Table 34

Perceived Central Focus of SB 191 Reforms (CF), by TAP

	Students	Teachers	Administrators	Policymakers	Total #s
Teachers	16%	59%	1%	24%	504
Administrators	43%	45%	2%	9%	53
Policymakers	38%	63%	0%	0%	8
Total	–	–	–	–	*565*

Value = 29.001, df = 6, Asymp. Sig (2-sided): .000

Note: Five cells had an expected count less than five, which may have impacted significance.

Table 35
Which group has primary responsibility for successful implementation of new teacher evaluations?, by TAP

	Teachers	Administrators	District Personnel	Policymakers	Other	Unsure	Total #s
Teachers	13%	56%	14%	4%	1%	13%	528
Administrators	4%	75%	18%	0%	0%	4%	55
Policymakers	0%	56%	22%	0%	11%	11%	9
Total	—	—	—	—	—	—	*592*

Value = 29.080, df = 10, Asymp. Sig (2-sided): .001

Note: Eight cells had an expected count less than five, which may have impacted significance.

perceptions disagreed more frequently than teachers' and policymakers' perceptions or building administrators' and policymakers' perceptions.

Open-ended response questions. In addition to the 40 Likert questions, survey participants were also invited to write an open response at the end of each set of questions. Not everyone choose to do so, but a number of participants did write comments (Table 36). Participants most frequently wrote to follow up on the questions

Table 36
Number of Responses to Open-Ended Survey Questions

Question Area	(Code)	Number of Participant Responses	Response Rate, within open-ended comments (%)
Current Teacher Evaluations	C-TE	131	25%
Ideal Teacher Evaluations	IdealTE	98	18%
Change	Change	77	14%
Education Reform	EdRef	111	21%
SB 191 Teacher Evaluations	SB191_TE	65	12%
SB191 in General	SB191	51	10%
Total		*533*	*100%*

about the current system and education reforms in general. The results of these responses will be explored in the qualitative findings.

Summary. The data analysis of survey results found that teachers, administrators and policymakers had differing perceptions about many elements of teacher evaluations, education reform, and SB 191. However, they collectively expressed a desire to participate in change and, in regards to the upcoming changes to Colorado's teacher evaluations, they believed that administrators will have the greatest role in the reform's success and that the reform was aimed at teachers. The subgroupings of teachers showed general agreement about perceptions regardless of subgrouping, with the occasional exception of teachers with a non-professional license and of a minority background. The subgroupings of administrators revealed a largely unified body of practitioners. The comparison of teachers', administrators' and policymakers' perceptions revealed frequent and statistically significant differences in perceptions between RSD teachers and administrators.

Qualitative Findings

Utilizing a Community Based Research conceptual approach of working with individuals who are members of the participant groups of interest in a study (Couto, 2003), interview questions were developed (Appendix E) for RSD teachers and administrators based on survey data. The objective of the interviews was to gain representative practitioners' insights into possible explanations for the data results. Because differences in perception between teachers and administrators were found in the survey analysis, interview questions were written to seek deeper understanding. The

interviews were transcribed by the researcher. The transcribed data were first coded through the software program ATLAS.ti, with a second and final hand-coding. The coding revealed a mix of topics, some with verbiage directly aligned with the question areas on the survey, such as "change," and others related to the question areas, such as "testing" and "tenure." The final categories were: "teacher evaluations" (specifically, "purpose," "tenure," and "time demands"), "testing;" "Senate Bill 191," and "teachers' voices in implementing school change." The theme of "fear" was mixed throughout the topics.

Survey participants had the opportunity to complete an open-ended response at the end of each section of questions on the survey. The six areas of questions in the survey were:

- RSD's current teacher evaluations
- Ideal teacher evaluations
- Education reform in general
- SB 191 teacher evaluations
- SB 191 in general
- Change

These answers were coded in the same manner as the interviews: first through ATLAS.ti and then by hand-coding. The categories for the open-ended response questions were again "teacher evaluations," "testing," "Senate Bill 191," and "teachers' voices in implementing school changes," with the additional topic of "non-classroom educators" and theme of "fear" arising throughout the four categories.

Interview analysis by group. Interviews with the four teachers and two administrators occurred on the phone over a period of three weeks using the semi-

structured interview format (see Table 4 for general descriptions of interviewees; see Appendix E for the list of interview questions).

Teachers. The teachers expressed differing perceptions about the role and nature of teacher evaluations in general. Shelley (all names pseudonyms; see Table 4) shared that teacher evaluations should be about "looking to build teachers and not to break them down." She advocated for evaluation systems to use a coaching model, one that would "make you [teachers] better." Jake did not believe teacher evaluations could be a used as a growing tool, for he viewed them as a tool for removing poor teachers rather than improving all teachers.

> I think almost no evaluation systems are designed to really distinguish among the top 25%. I think that they're set up to put people into three pretty broad categories: first, terrible, get out; second, you're good but you need some work or, third, hey, keep it up. And so, I honestly… I feel like most of the attention is going to go to the bottom two and not to the top.

Kate challenged the very concept of evaluations being on a traditional strong versus weak continuum, arguing that teachers do not view their abilities in such a black and white manner.

> I think when you talk to teachers about strong and weak, they're much more likely to say, "I'm really strong with kids who are struggling, I'm really good at sort of motivating them and sort of moving kids from the bottom to the middle." And other teachers say, "Gosh, I'm really strong with strong kids. You know, if they really want to learn, I can teach them." And that's what our data would show, actually, would be when we sort of break [it] down, that the teachers have different strengths.

The elimination of tenure has created an element of fear around the new teacher evaluations. Polly noted that the eradication of tenure was perceived as "scary" because it will require ongoing job evaluation. She further articulated:

> Teachers are feeling like every year we have to prove ourselves as talented at what we do. And there's something scary about that. Especially for older teachers who have gotten comfortable… There's something scary about being watched all the time to see if we're slipping or doing something wrong or not updating our curriculum. I feel like I hear a lot of fear in teachers.

Polly added a specific example of this anxiety:

> I just heard a teacher who has been teaching for 20 years and she's wondering if she's good at what she does because of this change, because of people coming into her room all the time. She's nervous that she thinks she's good at what she does, but she's not. And that's a very experienced teacher.

Polly also shared her belief that one of the reasons teachers were so concerned about this constant evaluation was because they regularly assessed themselves and their ability to reach each child in the classroom and, thus, this threat of job loss was putting on additional pressure.

Yet another concern teachers voiced about the new evaluations related to anticipated high time demands on teachers and administrators. The teacher evaluation system prior to SB 191 had teachers with professional licenses being evaluated, legally and on average, every five years. Under this approach, Jake believed that "administrators overestimate their own abilities" to evaluate a teacher through a few classroom observations and discussions. Under SB 191, teacher evaluations will happen annually for every teacher. This was anticipated to be highly demanding for both teachers and administrators. Kate shared that teachers already speak possessively of their time and that adding anything else to their schedules was "very distressing to teachers at this point." Both Kate and Polly questioned the logistics of how administrators would be able to evaluate all teachers on an annual basis under SB 191. Polly stated that "[o]ne of my

concerns about it right now is that if every teacher has to be evaluated every year, our administrators aren't going to be able to do anything else." Conversely, Jake thought this new system might empower administrators differently. Jake stated that "by having to evaluate their teachers half on student achievement, they've [administrators] just been given a *huge* gift;" on the other hand, Jake was quick to note that it was a "cursed gift" because of all the time needed to evaluate each teacher annually.

The teachers interviewed also expressed concerns, "a healthy amount of fear," about testing and the use of statewide assessments on teacher evaluations for many reasons. For Kate, one of the problems with statewide assessments was their lack of transparency. "The problem with the statewide data, from a teacher's point of view, is that it is essentially mystery data. We never really see the questions or how they were graded; all we see are very vague standards." This concern was corroborated by Polly who noted that "we don't trust the test, yet." Polly explained that "they present to us that they [test questions] will be thinking questions, high-level thinking rather than content-knowledge based, yet they will be graded by the computer. Right there is the disconnect in our minds." A second problem teachers had with statewide assessments was students' lack of accountability to the test. In Colorado, state assessments results have no bearing on a student's academic record and, therefore, students have little incentive to perform well. Jake noted that "the students have a very limited buy-in to it, other than what their intrinsic motivation might be to do well in the first place." Kate also articulated this concern, saying "there is no incentive for kids to do well." Students' lack of incentive would have large implications for teachers, though, as Kate revealed in a personal

example: “I had a student last year who was an honors student—he was very, very bright. He was angry that his mom had come to school during TCAP and so he got an unsatisfactory.” She went on to note that this one student’s result negatively affected her overall rating. A third concern teachers’ expressed with the use of statewide assessments on teacher evaluation was how data would be collected and used for teachers of non-state assessed subjects let alone non-classroom educators who will soon have a new evaluation as well. Kate relayed an experience she had with another teacher: “I talked to a PE teacher today, and he said to me, ‘So, Kate, I have to depend on you and your kids to give me a score because you’ll just talk about whole school scores? What if you’re no good? I can’t influence that!’” A fourth and final concern was that the use of statewide assessment data on teacher evaluations may create a negative, competitive environment among faculty in schools. Shelley articulated that the use of these data was “not even healthy” because it “does not support collaboration” and it “actually isolates people.”

The teachers interviewed expressed a range of views about the purpose of SB 191. For instance, Polly understood one of the objectives of SB 191 to be improving teachers’ craft, explaining that the purpose cannot be

> to get rid of teachers, because it gives you three years to stink. So I think if it was to get rid of teachers, it would be one or two, right? This is three years to grow: the intent is to grow teachers. There are definitely some best practice things that are learnable, teachable, that can really change a teacher’s success, so I really feel like the intent is there.

Kate’s perception of SB 191 was that the bill’s “purpose is to improve teacher effectiveness, which improves student learning; not to make a teacher less comfortable in his or her job, or to put more onus on a teachers than they already have.” Jake understood

the bill differently, stating his belief that the bill would be used to siphon off weak educators rather than to grow teachers. “Senate Bill 191 is about punishing teachers who don’t get students to grow.”

Despite teachers’ uncertainty about the purpose of the bill, there was some recognition that it may improve student learning. Jake articulated his view: “I can’t in good conscience say that SB 191 won’t push people towards trying to have more student achievement; I think that’s going to be one of the effects, that it will push.” Polly shared this belief, stating:

> I know that there are teachers who assign things and then sit at their desk and shop for shoes. I know they’re there. If this inspires them to stand up and teach, then I feel like that alone is going to impact student achievement. Now, if they’re only standing up and teaching on their show days, then no. I’m hoping that it does actually impact the do-nothings that are in our profession; I do think that is one of the intents behind it.

Shelley also noted that in her school there are “a couple teachers, a low percentage of teachers, that aren’t doing their jobs and that’s to say the least,” who may work differently because of the requirements of the new teacher evaluations. Yet, she also shared her belief that this would be minimal because

> Overall, teachers get involved with education because of what they want students to learn and they want to do well for their students and they want to have their students to grow and they invest a lot in children in general.

However, the means of getting teachers to make this shift were perceived to be fear-based. As Jake noted, “it’s just a question of how... other than their personal fears.” Kate believed a fear-based approach would be problematic and ineffective; she referenced the author Daniel Pink’s work on motivation, noting that Pink

> makes a very good argument that the more complex the task, the less likely incentive-based programs work. I think a lot of teachers are sort of aware of that, at least in a general sense. It's my feeling, and this may be Pollyannaish, that many teachers really desire to be excellent, and telling them that there are strong and weak among you and that this rubric is going to help you get better... I don't think it necessarily creates confidence in them that that will work, in terms of kids actually getting better.

The teachers interviewed believed that RSD teachers' voices were being only sporadically included in implementation of new evaluation system. The interviewees each noted that teachers in their schools were generally aware that there would be a new evaluation system next year, but that most were not directly involved and fewer still were familiar with the law. Polly noted that she was the only one among her peers who had read the legislation. Polly, Jake, and Kate all shared that teachers from their school, "teachers that are loved and respected," were involved in the pilot process and that this was interpreted as being positive because then teachers could ask their colleagues and administrators about the new system. Polly stated that "the doors have been open, but informally" for all teachers to participate in the implementation of the new evaluations. Most teachers, Polly believed, were "just hearing about it [the new evaluations], it is so passive for the teachers, most of the teachers are like, 'oh, here it comes.'" Jake also articulated that few teachers were directly involved, stating "in most schools teachers aren't asked to participate, they're told what to do." Yet, Jake noted, this lack of involvement could be challenging in schools because "[t]eachers take it [evaluations] very, very, very personally." Kate also voiced concern about the lack of teachers' regular involvement and the implications of such, stating that teachers "don't believe that they

have any influence over this process, really, and therefore they feel that improvement is not going to happen in schools."

Each teacher shared that the best way to transform SB 191's teacher evaluations into a meaningful and lasting reform in their school would be to include teachers' voices in the changes as much as possible. Jake stated that his principal has a pattern of including teachers in piloting major changes, which had a positive impact on the teachers. Shelley articulated that "[t]he right way [to introduce change] is to build readiness, so to prepare for what's next. And the wrong way is to jump in without teachers' understanding what is fully being expected." Polly expressed similar sentiments, saying she thought her colleagues would do best by hearing "we care about what you think and we want your voice" from school administrators. She also articulated that administrators should directly "address what you know the teachers' fears are." Kate also shared hope for a collective approach: "I think the right way, and the way that we're trying to do, is to talk about it as a team approach."

Administrators. The two administrators interviewed, Andrew and Anna, focused on the changes SB 191 was bringing to teacher evaluations. Overall, both supported the changes. Andrew stated the legislation was "well-intended" and Anna articulated "I agree with the law and I'm very supportive of SB 191."

The eradication of tenure in connection with the new teacher evaluations, though, drew mixed responses. Anna expressed her belief that this represented a meaningful shift in practice.

> [W]e've been living in a system where, after three years, teachers, you know, gain tenure. Now we're moving to a system where if you have two consecutive years of underperforming performance, and it's a whole other game. This is changing the playing field *completely* in education.

Anna noted that this shift may be making teachers uncomfortable, at least initially, with the new evaluations. She added that there are many "people with high emotions [who] are currently in fear of losing their jobs." Yet, this very shift in the evaluations was what Anna thought would allow administrators to help teachers improve the craft. "[A]dministrators are saying, 'well now we have an opportunity to really help teachers grow with a document that will really help them grow.'"

Andrew's beliefs differed. Andrew believed one of the objectives of SB 191 was "to help administrators remove ineffective teachers who were non-probationary," but he did not anticipate this would happen in practice. Andrew believed that the replacement of tenure by the new system—two consecutive years of ineffective ratings leading a teacher to a year of probationary status before the potential for dismissal—would be actually be "more restrictive" than the previous practice and turn into a "game that teachers will understand how to play."

The student assessment growth data component of the new teacher evaluations worried the administrators for two reasons. First, like teachers, they expressed concern about student buy-in. Andrew empathized that teachers were concerned about their students not performing well on the statewide assessments. "I think there's fear that no matter how hard they [teachers] work, there are going to be students in their class who are going to struggle [on statewide assessments]." He further noted that, "on a teacher

level," how students perform on assessments is "very personal." Anna shared a related concern that teachers working with certain groups of students, particularly Title One students such as those with increased mobility, free and reduced lunch, and English Language Learners, would be less likely to see their students make significant progress on statewide assessments and that this could harm such teachers' evaluations. She was quick to add that evaluators would not view these students' poor performance as indicative of their inability to learn, or teachers being unable to teach them, but would consider factors external to school that could make it more difficult for these students to perform well on this type of assessment. The administrators' second concern with the use of statewide assessment data on teacher evaluation was the timeliness of results, or the lack thereof. Historically, Colorado State Assessment Performance (CSAP) results arrived at schools in June. New statewide assessments will be coming out in the next year or two, but if the results were still to arrive after the school year was over, Andrew wondered: "How is that really going to work? Is the evaluation not complete until the following school year?" It was unclear to Andrew how assessment data would be usable for teacher evaluations under the statewide assessments' former timeline.

The linking of student assessment data to teacher evaluations did have some positive elements for administrators. Administrators, Andrew shared, were "used to having data be part of our evaluation process," so this addition was comfortable for him Moreover, Andrew noted that previously he had struggled with getting teachers to understand the importance of state testing, and so he actually welcomed this change.

> One of the challenges that I have, that any principal probably has, is making sure that there is ownership of those scores at the teacher level. I think that an administrator welcomes this idea because it will help us with teachers having buy-in in students' performance on standardized tests because it's also going to affect them directly.

Anna also expressed general comfort, in the long run, with the connection between statewide assessment data and teacher evaluations. She stated her belief that the alignment of statewide assessments to both the common core, as will soon be done, and to teacher evaluations will be a "win-win for kids."

The need to conduct annual evaluations of all teachers was anticipated to be challenging for administrators due to its high demand on administrators' time. Andrew articulated his concern that the amount of time he will have to spend on meetings and paperwork for evaluations will reduce the amount of time he can spend on other elements of running his school. Andrew anticipated that the time demands would be so great that he doubted the bill was "really going to help students because now the amount of time that administrators are going to have to dedicate to this system is going to be huge." Andrew also noted that his faculty had similar concerns.

> The common thing that I got from that [in-service] day was that there was a concern around how the administrators would be able to do all those evaluations and still be able to do all the work that teachers are accustomed to us doing.

The new administrator evaluations are also expected to be more time-consuming than the previous one, potentially compounding the time challenge for administrators. Additionally, Andrew anticipated that teachers would "probably question" whether or not administrators can conduct thoughtful and thorough evaluations of all teachers on an annual basis.

The administrators' interviewed each expressed a desire to have teachers included in the implementation of the new teacher evaluations and used the phrase "transparent" many times to explain the process at their respective schools. Anna stated that there was "no hidden agenda" and all information she was presenting to her faculty members was available in handouts and online for additional reading later. Both Anna and Andrew talked about the benefits of the pilot program during the 2012-2013 school year, namely that the process required one teacher from each school in the district to be directly involved. Anna noted that these teachers have already been trained in the new evaluation, which "helps teachers feel like they are a part of the process." Andrew shared that he had devoted an entire in-service to explaining the upcoming teacher evaluation changes and that, in regard to feedback, "100% have been given that opportunity." At the same time, Andrew stated that the percentage of teachers who had been actively involved in the pilot process and might be familiar enough with the new evaluations to ask tough questions, versus those who had only cursory knowledge of the new evaluations, was less than five percent. For this reason, he said: "I would say that that probably is fair to say that the majority of the teachers feel like they're not going to have a direct impact."

The administrators believed that part of their job as administrators was to lead the changes accompanying the implementation efforts. "It is my primary responsibility [as principal] to roll that out. Absolutely, I agree with that." Both also expressed a desire for the implementation of the new evaluations to be one of "shared ownership" between teachers and administrators. Andrew commented that "[t]he way we've unrolled it, feels like a real partnership." Andrew also noted that "[i]f we're in this together, then hopefully

that will help with the transition to the new system." In speaking about RSD in general, Anna stated, "I can speak for RSD and how we're integrating our association, different employee groups and really making it a partnership. And putting the ownership back on the teachers, and really front-loading it really well." Anna also noted that, despite the focus on administrators and teachers partnering to implement the new evaluations, she believed there were still teachers uncomfortable with the situation: "it's that fear, it's that unknown, it's that change." Additionally, she noted, that teachers are "very emotional. I would say this is a very emotional topic," and that this contributed to that fear of the unknown elements of the new teacher evaluations. But she also stated that all teachers should be becoming familiar with the new evaluations by this point.

> I would disagree with the fact that they don't have access to what it is because… we have given every employee access to what it's about. There's no surprise. We have been talking about this for 18 months that it's been coming.

Responses to open-ended survey questions. In addition to the six interviews, qualitative data for this study also came from survey participants' responses to the open-ended questions that followed each area of questioning on the survey (for details about responses to the questions, see Table 31).

Teacher Evaluations. The participants' perceptions of teacher evaluations showed that, collectively, they wanted evaluations to be used for growth but did not always see this happen in practice. One participant wrote that evaluations were "meaningful" and allowed for "cooperative learning" for teachers and administrators. Another stated that evaluations were "helpful and thought provoking." Yet another asserted that "[t]eachers want feedback and to become better." Still another wrote of the benefits of constructive

criticism: “Whenever I was given critique from my principal during evaluations, I found them very helpful and always implemented those changes. I was lucky to have a good principal who made me a better teacher.” Other participants, though, questioned the ability of teacher evaluations to improve practice. One wrote, “I do not think it is used as a tool for teacher improvement. It tends to be more summative and snapshot specific.” Many participants expressed concern over the use of teacher evaluations to improve practice because of the role administrators played in the teacher evaluations and the process itself.

A number of participants critiqued the teacher evaluation process for being too dependent upon administrators’ strengths and weaknesses as evaluators. Participants expressed this concern in many ways: “[t]he success of this process depends on which administrator is doing the evaluations;” “growth and feedback is dependent upon WHO” is evaluating you;” “[i]t is an uneven experience that varies from administrator to administrator.” Another participant shared that “evaluations are often about opinion with very little subjective data.” Yet another related “[s]ome evaluators are better than others, so there is discrepancy in how well any evaluation tool is implemented.” One participant shared an overview of her experiences:

> I have had three evaluators. My first was amazing. I had meaningful feedback, got great suggestions and I really felt like I grew as a teacher. My last two were not nearly as meaningful. My feedback was generic and more ‘copy paste’ than individual feedback. I do not feel I grew as a teacher from the last two evaluations.

Yet another participant echoed the same concerns about the variability of administrators, while also noting teachers’ role in gleaning meaning from evaluations, stating

> The teacher evaluation process may provide teachers with valuable feedback, but it is dependent on the effectiveness and thoroughness of the evaluator and the teacher's reflectiveness.

Many participants also challenged the ability of teacher evaluations to assess teachers thoughtfully and thoroughly when teachers often were observed only for a few pre-determined lessons. The most frequent critiques were that administrators saw merely "snapshots" of a teacher's ability and that most classroom observations were "dog and pony shows." One participant wrote:

> Two observations of teachers during an evaluation period simply promotes a "dog and pony show" that can be put on by young teachers or teachers who may not be strong in the classroom. Anyone can fool anyone for two 45 minute blocks of time.

A number of participants expressed concern that the brief nature of observations masked weak teachers' true abilities and, therefore, the participants advocated for more frequent and unexpected observations.

> Most teachers that are ineffective are ineffective because they are inconsistent and/or lazy. They can put on a good show and teach a good lesson when they know they are being evaluated. Unannounced evaluations would fix most of the issues, I believe.

Another wrote "I do, however, feel that we, teachers, should not be the ones to choose the class period that will be evaluated. I would prefer that my evaluator observes me teaching under unplanned conditions." Participants also noted that the brevity of observations was due to administrators' time constraints, rather than to administrators' lacking that ability or desire to conduct more thorough observations. One participant noted that evaluations were "generally poorly administrated because of the lack of time that administrators have

to do the teacher evaluation." Another wrote that it would be nice if an administrator could spend a week with a teacher "[b]ut that requires a lot of time which isn't there."

Participants expressed reservations about the new teacher evaluations because they perceived it would place even higher demands on administrators' time. "The law puts a huge burden on administrators," one participant wrote. Another wrote that "[t]here is not enough time to evaluate every teacher. Administrators are out of their buildings enough, now all they would do is paper work when in the building." Yet another shared "I am most concerned about how our administrators will be able to implement this evaluation process on top of all of their current duties." Still another participant voiced this concern:

> Building Administrators should have primary responsibility for implementation of the new teacher evaluation system but many administrators don't have time to invest in its 'goodness' so that we can feel that it will be beneficial for all involved. Oftentimes, I feel that administrators are just doing what district has told them to do. It has been pushed on them, so they just give it to us and say please do. Many department chairs don't understand and thus can't explain the reasoning so it becomes just a task that is required by district and the reasoning is muddled somewhere in the "please get this done."

Administrators' knowledge of teaching practices and ability to provide meaningful feedback were also questioned. "Often times the administrator has had little to no actual classroom teaching experience. This makes it difficult to have a conversation about teaching in an open, honest, and informative manner." This concern arose repeatedly among educators who were not regular classroom teachers. One wrote:

> I am a special [needs] teacher and do not feel that I ever get any meaningful feedback about my specialty area. I wish there was an evaluation process that would address what I do specifically and give me feedback that pertains to my

> specialty. I dread filling out the form each year that I am evaluated because it's an absolute joke!

Another participant shared specific examples of how an administrator evaluating him/her was unable to help improve practices:

> I am evaluated by an administrator with no special education background (much less any knowledge of the speech-language field). I would love to improve my skills as a speech-language pathologist–but my principal will never be able to give me new ideas/suggestions on how to reduce a student's stuttering, program an augmentative communication device, shape sounds around a cleft palate, etc. So how does this process help me improve at my job??

In contrast to the participants who voiced frustration with administrators' inability to provide knowing, constructive feedback, a handful of participants commend their school's peer-based teacher evaluation practice for providing just that. One participant applauded the process, stating:

> I have been able to participate in peer evaluation. I felt this was very encouraging and has helped me improve some of my teaching strategies. As teachers, we don't get the time to observe other teachers, and I found this enlightening, informational and [it] encouraged me to do some things differently. I found it to be more beneficial than being observed by an administrator.

Other participants share similar positive experiences. One noted:

> I received more thoughtful feedback from peers than from principals, who are so stressed with the number of observations they must do [that] they find it difficult to provide me with the in-depth feedback they would like to give.

Another commented that he/she "probably worked harder" on the goals set with a peer evaluator than any lesson plans he/she might have developed for a more a traditional administrator evaluation.

Many participants connected the elimination of teacher tenure, arising from SB 191, to the creation of a negative work environment. One participant noted that, under the

outgoing teacher evaluation system, evaluations were "not directly linked to teacher performance or student learning unless the teacher is targeted for termination." Another participant wrote, "[d]o not make it seem as if we are all going to lose our jobs with a bad review. We need to understand the review process and work out the kinks before we begin to attach teacher retention to the new system." Yet another participant commented that "I am not afraid of the new evaluation system, but many of my peers seem to be." Still another shared that the "stress of the environment for a lot of teachers is problematic as it is breeding a culture of fear. Improving practice is not going to change out of fear."

Testing. Many participants commented on some element of the role of testing in connection with the new teacher evaluations. A handful of participants applauded this addition. "I think linking performance to data is an effective practice." Others challenged the idea of including any type of student assessment data as part of teachers' evaluations. One participant wrote that tests were not an acceptable measure of teachers' abilities because a test "is testing the student's abilities to take a test, not the teacher's ability to teach the student." Others questioned the use of tests for pedagogical purposes, such as one social studies teacher who wrote "I believe that testable factual knowledge should usually take a backseat (though it certainly is still important) to critical thinking development." Another wrote "many state-made tests don't measure 21st century skills, which is what many districts/schools are more focused on implementing."

Of the participants who commented on the inclusion of student learning growth data in teacher evaluations, most presumed that these student data would be taken from statewide assessments; this elicited myriad criticisms. One participant wrote that

"[l]inking teacher effectiveness to corporate standardized assessment data is insane." Another commented that "[s]tate department of education testing items and procedures fail to accurately assess student learning." Yet another expressed concern that "common assessments are just ridiculous" for they directly contradict the idea of "individual instruction, students as individual learners" that the district espoused. Another participant added that "[t]eachers should only be evaluated on the procedures and practice within their control," which the participant noted did not include state designed assessments. Other participants wrote that multiple assessments should be used instead of just one. The fact that a statewide assessment for students (to replace CSAPs) was still being developed at the time of the survey drew these questions:

> What assessment will be used? Who will decide what assessment will be used? Will it be one assessment, or more than one? How will the poor scores of students with circumstances out of my control impact my evaluation? All of these questions are unanswered. It is unacceptable that I will be evaluated on measures that are still unknown.

Yet another concern around testing was that it was being appropriated to measure teacher effectiveness instead of student learning. Following the statewide assessment tool's upcoming unveiling, participants expressed concern that it be "normed and validated" for determining educator effectiveness as well as student knowledge, if it would be used for these dual purposes. Many participants stated concerns that statewide assessments being used to assess educator effectiveness represented a "misuse" of the assessment. In related fashion, a participant wrote that "data should drive instruction, not performance."

The timing of state assessment results, historically delivered to Colorado schools in June, was also perceived as problematic. One participant wrote that "[t]eachers should be able to adjust teaching methods and modalities from the data," which would be possible only if the results were provided during the same school year in which assessments were given. Another participant stated that "teachers are unable to access data in a timely fashion and use it to further student growth." Yet another echoed these same concerns and provided a suggestion, stating:

> State data is delayed and offers the teacher no opportunity to evaluate their students and target instruction for the students who are not proficient. Periodic benchmarks would be more valuable if there were immediate grades available.

Many participants expressed concern about the role of students in the testing process, both in terms of students' desire to perform well on statewide assessments and students' backgrounds impacting performance on such assessments. Multiple participants shared their beliefs that "students have no buy in" to statewide assessments. Or, phrased different, "[s]tudents have absolutely NO responsibility to do well" (capitals in original) for, as yet another participant wrote, students perceived such statewide assessments to be "irrelevant… and have no effect on them." One participant shared this anecdote:

> I personally have a friend who told her five children that she does not care how they do on these tests. Her children fill in all the C's, and write two sentences where essays were required. I do not believe this is an isolated case.

A number of participants suggested that state assessments have some bearing on students' academic records if this would influence teachers' evaluations. "Both teacher and students should have the same desire to have the students do well on the test," one participant wrote.

Students' backgrounds, and how they impacted performance on assessments, drew many comments. A few participants focused on students' familial backgrounds. One participant wrote succinctly, "as we all know, CSAP data correlates directly with SES status of student." Another contributed that "[a]ny standardized test you use is biased to the wealthy and the White. Period. That makes inaccurate data and the whole house of cards falls from there." One participant shared this example:

> [A] few years ago there was a question on CSAP about a yacht—we live in landlocked Colorado—many students had no clue that it was a big fancy boat. As a teacher, I should not be penalized because the test is not culturally relevant to my students.

Still others expressed concern about factors related to a student's well-being and ability to perform well on an assessment on any given day. Wrote one:

> Tests given once a year on one or a few days does not take into account other factors–health of the child, overall attendance of the child, how long the student has been enrolled in a course, level of course student is currently enrolled, etc.

Another participant shared similar concerns about the use of a single assessment to judge both students' learning and, soon, teachers' effectiveness:

> There are too many variables in the lives of our students and one day could be a great day to test while another is not. Many of our student's do not come to school with food in their tummies, a good night of sleep, and they may have just experienced a traumatic event that morning that will distract them during the whole testing period. There are too many outside factors that play a HUGE role in the lives of our students and it would not be an accurate measure to evaluate a teacher on classroom assessment data. (capitals in original)

Yet another participant shared that, due to "difficult home situations and a lack of parent support," there was a "large turnover of students" in many schools, which would make it "impossible to really assess how the teacher is performing" based on students' scores.

Teachers of students with special needs shared specific concerns about using statewide assessments to evaluate students' learning. One participant shared two specific examples of how using statewide assessments with his/her particular students could be challenging. First, he/she wrote that, as a teacher

> in a setting where students are generally behind socially, academically or both[,] we have to look for alternative assessments; i.e., if we are trying to develop work skills, the fact that we get the student to take a shower once a week instead of once a month is a big win. How will CDE take that into consideration?

Second, this teacher shared how, when he/she once "got" an eleventh grade student who arrived with a fourth grade reading level and advanced to an eighth grade reading level, he/she considered this a "win-win;" yet that student would still not perform well on statewide assessments. Another participant wrote that there should be

> a specific and reasonable plan for those teachers that work with students who have needs vastly different from the 'typical' student and especially those teachers of students that take alternative tests. In other words, teachers should not be punished for their students' disabilities and limitations.

Yet another participant wrote that teachers of students with learning differences could not have their students' statewide data applied to their evaluations for "students with a TRUE learning disability often lack the ability to consistently APPLY their skill set to testing content." (capitals in original)

Still another concern related to testing was that the focus on these statewide assessments detracted from overall learning. One art teacher shared her perspective:

> I know how powerful art is on brain development especially for my dyslexic students and those with learning disabilities. I do not believe that many people, the CSAP test [writers] included, understand the important connection to developing a child's brain development. CSAP data does not accurately reflect a child who has dyslexia or L.D.'s knowledge.

Another participant wrote that "[a]uthentic teaching (and learning) is jeopardized when the goal focuses on the attainment of higher test scores." Similarly, one participant wrote that past statewide assessments have counteracted his/her way of teaching, for students were "taught to revise and edit using resources but then they can't use those resources on the test." Another stated that the CSAP was "such a poor assessment of a students' math skill/learning/aptitude that one must interpret CSAP results very carefully."

The theme of fear was scattered through participants' writings about testing, but particularly in regards to its potential impact on school culture and teacher collegiality, on teachers' willingness to work with disadvantaged and special needs students, and on teachers' lack of control over test-day conditions. One participant wrote that "[h]igh stakes testing attached to teacher performance breeds cheating and resentment among teachers." Another revealed perceptions about the connection between teacher evaluations and testing, saying "let's help teachers who are struggling rather than punishing them immediately through the use of TCAP scores." A handful of participants raised questions about teachers' future willingness to work with students who may not make measurable growth on assessments because of the anticipated connection between statewide assessments and teacher evaluations. One participant stated

> I don't think this has been thought out very carefully. Some teachers end up getting the lower tougher kids year after year because they are better at meeting certain student needs and how will this play out?

Another participant asked "[w]ill the additional scrutiny cause teachers to not want special education students included in their roster?" A special needs teacher wrote:

> now under the new program, my performance will be based on the academic growth (as measured on standardized tests) of severely impacted children with special needs–when by definition, they qualified for special education services because they do not make adequate growth EVEN WITH ongoing specialized instruction. Completely frustrating and bureaucratic… (capitals in original)

Teachers' inability to control for numerous factors outside of their control also raised concerns. One participant wondered if "allowances should be made for variables that are outside of the teacher's control," such as "motivation, mental health, family environment and truancy," since all these factors could directly impact a student's assessment yet may not be directly connected with a teacher's teaching abilities.

Senate Bill 191. The participants showed a full range of degrees of support for SB 191 as a reform. Some participants endorsed the measure fully, with one writing "I believe focusing on student growth is the crux of our jobs and shouldn't change too much for effective, committed teachers." Another stated "I whole heartedly agree with the purpose and intention behind the new teacher evaluation SB 191." Yet this same individual acknowledged doubts about the bill's ability to succeed.

> Like anything else in education reform, the time and money to actually implement a significant paradigm shift in a meaningful and purposeful way for teachers, building and ultimately kids, falls well short. [To evaluate] every teacher, every year, with one building administrator is unfair and dumbs down the purpose of the new teacher evaluation–I fear that it will come down to pushing paper versus real cognitive coaching around teaching effectively due to the lack of funding for administrators (more people) to evaluate everyone with fidelity.

Other participants shared this type of skepticism. One commented that Colorado needed to "increase funding to coincide with the new policies. We are grossly underfunded." Another wrote that evaluation "reform is necessary, but this bill has been completely rushed, mishandled, and focuses an unfair and disturbingly negative lens on

teachers, and puts an unfair burden on administrators." One participant expressed doom when writing about SB 191.

> With the implementation of Race to the Top and SB 191, we are witnessing the systematic privatization, corporatization, and financialization of the public education system. Corporations, the testing industry, and state and federal lawmakers are the only ones who stand to benefit from the implementation of standardized tests and the Common Core standards. Teachers and students will ultimately suffer from these self-defeating ed policies. I fear for the future of public education in America.

Yet another participant shared concerns that went beyond SB 191's reform efforts, stating that reforms led by the

> [g]overnment whether at the state or federal level often lead to unintended consequences (like increases in paperwork, inefficiencies in getting resources to the right people, etc.) and have little result on changing education… The U.S. outspends all of these [other] countries on education reform and we wind up empty handed. When government places the burden of societal change on the teacher or the educational system at large, and not on the primary influences of a child's life like their family, the results are inconsequential.

Still other participants, though, expressed their concern that SB 191 did not go far enough in its reform efforts.

> Real reform means real change in the way we "do school." The world has changed, and the structure of school day and year has not kept up. Raising the bar is very good, but woefully inadequate when it is sum total of the "reform."

Many participants shared that they understood SB 191 to have been labeled as a law about students and student learning, but they believed in practice it was about teachers. "I believe that the original intent of the bill was for the student but was lost somewhere in the process." Another wrote that, "[y]es, I know the central focus of the bill is supposed to be students, but I don't think it really is." Yet another expressed similar

views, stating the "bill is supposed to be about students. Teachers aren't fooled." This individual expanded on his/her statement, adding:

> We know we'll continue to have to prove we can do more and more for less and less. I often wonder if the intention is to make teaching in America about a 5 year profession. That way no one can accrue a higher rate of pay, let alone realize their full retirement, and that would save money. Further, less and less experience on the job would confirm the ever present belief teachers are incapable. My informal "polls" suggest, if given another opportunity (with at least the same rate of income), every teacher I know would find other work.

This sentiment was expressed by another participant, who wrote:

> Senate Bill 191 will stress out teachers even more than they already are. I am counseling young people interested in the teaching profession to find a solid private school and stay away from public education. It is the most loving advice I can give them.

Yet another participant expressed comfort with the bill's focus, writing:

> I say that teachers are the main focus because the bill is all about teacher effectiveness. Teachers are measured by the success of their students. While the goal is always to improve students, and make sure they are growing academically, SB 10-191 revolves around how to adequately rate the effectiveness of educators.

Many participants welcomed higher accountability for teachers as was anticipated will happen via SB 191's new evaluations while, at the same time, questioning the role of students and parents in the learning process. This was apparent in many participants' comments about students having the necessary buy-in to try hard on statewide assessments. The following comment was representative of many:

> Policymakers would have you believe that every student arrives at school fed, properly dressed, washed, and eager to learn–if only the teacher doesn't let him or her down. That's certainly not the case in our public schools–not every student is interested in the process of learning, and no amount of legislation is going to change that. The teacher evaluation system could use some augmentation, but it is not a cure-all for the failure of some students to take advantage of what their public schools are offering them. The number one factor in a student's

> achievement is the student himself, but some policymakers would have us believe otherwise.

Yet another participant asserted his/her belief about all educators trying to improve student learning, despite public perceptions to the contrary. "I believe that all teachers are in the profession to make students and their learning excel. All policies and their implementation should also assume that."

The role of policymakers in setting the agenda for education reform in schools drew mixed responses. Some participants supported the joint efforts of teachers, administrators, and policymakers. "It [reforms] should be a shared decision making process between proven leaders in administration, teaching, and policymakers who are willing to go into schools to understand education." Another participant shared similar sentiments: "I do not believe that the implementation of reforms should come from one single area of professionals. The best reforms will come out of those that include teachers, administrators, and policymakers." Yet many other participants expressed concern and discontent that policymakers were involved. One asserted that "the one thing that is common amongst the professionals in the field is that policymakers have no business determining what is best for our profession." Yet another stated "[p]olicymakers leading implementation of educational reform is like TV producers leading implementation of medical practice reform." Still another wrote, "[p]olicymakers have no clue about the classroom and what is best for students. It is about politics. I would liken this reform method to a layman/atheist teaching religion." One participant asserted: "[p]olicymakers making change is one of our LARGEST problems!!!"

Teachers' voices and implementing changes in schools. Participants expressed varying perceptions about whether or not teachers were being included in the implementation process of the new teacher evaluations. Some participants felt their administrators were bringing teachers into the process. One wrote that "[o]ur school administrators do a really great job of involving staff in change and empowering staff to share opinions/feelings." Another shared "[i]n general, I feel our administration does a good job of including staff in policy changes." But many more participants wrote that they did not feel included in the change process for the new evaluations. "Our admins speak down to us," wrote one participant. Another stated that he/she felt "like a puppet" when it came to school changes. Yet another participant wrote "teachers had very little input into the new evaluation system." Still others wrote: "[e]ducators working at schools have received close to ZERO information about how SB 191 is to be implemented" (capitals in original); "[s]adly, I know nothing about this bill;" "I don't really have any information other than the information in the newspaper;" and "I have learned absolutely nothing about SB 191 from my school district." One participant wrote that "there was an increased feeling of fear" due to the "lack of information that we have" about the new teacher evaluations. Teachers' lack of knowledge about the new evaluations led one participant to write that "[t]eachers do not have a voice to identify which aspects of the new process are unreasonable."

Participants' perceptions about the exclusion of teachers from the implementation of the new teacher evaluations contrasted with participants' expressed desire to be involved in changes in their schools. One participant succinctly explained the desire to be

included: "[w]e're humans… we like to be in the loop." Other participants focused on the connection between inclusion and one's willingness to work for the changes. "Ownership is necessary in getting anyone to adapt to change." Another stated "[a]s a learner, and it should be the expectation that all teachers are learners, I navigate change best when given the opportunity to take some ownership of the change." Yet another wrote:

> I think that the more ownership teachers have in a decision, the more likely they are to value the new policy. With whole staff buy-in comes school unity and alignment, which means that the change will be more effective in the long run.

One teacher shared:

> Teachers are generally loyal, over-achieving, hard-working employees. We are motivated by many things, but we learn not to depend upon appreciation or recognition from anyone in order to do our jobs well. However, I believe most will admit that being valued certainly encourages our best dispositions and actually works to foster ingenuity as well as create positive energy.

Other participants approached teachers' desire to be included in changes at their school from a different angle. One wrote that "[w]hen people are told what they have to do rather than being involved in creating the policy, it creates resentment and a feeling that one is not in control ([when] all people would like to feel control over their lives)." Another noted that not taking teachers' "voice[s] into account" would "deprofessionalize the profession." Similarly, one participant noted that precisely how changes were introduced in a school also mattered.

> How new policy is introduced and implemented is of the utmost importance. Teachers need to be included in this as they are the professionals who will have to work within the confines of the new policies implemented.

A related concern held by participants was the frequency with which schools experienced changes. One participant wrote that "[i]n the school system the policies change at work too often. There is less and less time to sit with a policy before another is implemented." Another noted that "[p]olicies can change several times with-in the same school year. This can be very frustrating." Other participants focused on the need for teachers to have time to process change, in order for the change to be followed with fidelity. "We also need time and resources to collaborate in authentic ways that propel the kind of lasting and positive changes the students, teachers, administrators, taxpayers, district, state and nation wants to see." Yet another wrote that teachers "need time to process changes and should receive professional development (and on-going coaching) in order to implement changes thoughtfully." Still another participant focused on the emotional nature of change, stating that "[m]ost transformations are going to be uncomfortable for a time. But, with time, we adjust. Gradual change with complete transparency to everyone involved is best."

As much as participants expressed a desire for teachers to be involved in changes at their school, they voiced concern about policymakers being involved in school changes—or at least policymakers doing so without direct input from teachers and administrators. One participant shared his/her concern with policymakers making changes in schools in this way, "I think policy is created with a good intent but it's being created by people not actually carrying out the implementation. THERE IS A DIFFERENCE ON THESE TWO POSITIONS!!" (capitals in original). Some participants expressed concern about policymakers being involved in school changes

when few of them have spent time in schools or classrooms. Many participants shared their hope for teachers, administrators, and policymakers to work collaboratively to design reforms for schools. “If all three groups can’t come together to figure out the right, meaningful things to do and how to do them, then we won’t improve all that much.” Another participant wrote:

> I don't think any one group should unilaterally decide what and how reforms are created; it should be a collaborative process for any change to be effective. Each of these stakeholders has a different perception and we need all the voices to share in order for reform to work.

Still other participants expressed concern that reforms were focused on political ends, with one participant writing that teachers and administrators’ were frequently discounted “in favor of dated surveys, research, and political agendas.” The primary reason one participant wanted teachers involved, though, was very direct: “[i]f teachers were included in policy changes there would be more meaningful changes in the system.”

Summary. Data from the six interviews (four teachers and two administrators) and the survey’s six open-ended response questions provided insight into how teachers and administrators in the RSD felt about teacher evaluations, education reform, Senate Bill 191, and school change. The topics of testing, non-classroom teachers, and teacher voice emerged, as did the theme of fear. Teachers and administrators had different perceptions about their own roles in the implementation process of SB 191.

Chapter Five: Discussion

Introduction

Despite the proliferation of reform efforts in America's public schools, few have made lasting positive impacts in recent decades (Sarason, 1990; Tyack & Cuban, 1995; Cuban, 1998; Tyack & Tobin, 1993; Hargreaves & Shirley, 2008; Berliner, 2005a). Colorado's Senate Bill 10-191: The Great Teacher and Leader Act is one of many laws nationwide currently focusing on improving student learning through the revamping of teacher and principal evaluations and the elimination of teacher tenure. The objective of this mixed method study was to understand how teachers, administrators, and policymakers perceive their respective roles in the implementation of SB 191's new teacher evaluations. The secondary questions for the study were: Within these three stakeholder groups, do responses vary by demographic variables? How do the responses from these three stakeholder groups compare?

This fifth and final chapter will analyze potential meanings behind the data. To do so, the quantitative and qualitative data will be combined. First, the quantitative data will be reviewed, focusing on possible implications of results from the comparison of teachers, administrators, and policymakers responses to the 40 survey questions. Second, the qualitative data will be added to the analysis of the quantitative data, as it was gathered expressly to help explain survey results. Following the interpretation of the

quantitative and qualitative data, there will be recommendations for RSD, additional limitations acknowledged for this study, areas for future research, and concluding ideas.

Interpretation of Data Findings

The survey data for teachers', administrators' and policymakers' collective perspectives on individual questions and groupings of questions was relatively straightforward. The data illustrated that teachers, generally regardless of subgrouping, were skeptical of education reforms and SB 191's ability to make meaningful changes to teacher practices or student learning. Administrators, regardless of subgrouping, were largely optimistic about education reforms succeeding in public education and about improvements in teacher practice and student learning resulting from SB 191's new teacher evaluations. Policymakers, as well as could be determined from the few responses, landed somewhere in between, aligning alternatively with teachers' and administrators' views. The juxtaposition of these three stakeholders' responses merits additional consideration.

As stated in Chapter Four, of the survey questions examined via Pearson chi-square analysis of teacher, administrator, and policymaker (TAP) responses, 19 questions had statistically significant differences necessitating additional analysis by a one-way ANOVA with Scheffé post-hoc tests (see Table 27). The ANOVA results illustrated that teachers and administrators had statistically significantly different results 15 times, while teachers and policymakers had statistically significantly different results one time, as did administrators and policymakers. Additionally, for the 10 questions asked of just teachers and administrators (see Table 27), seven resulted in statistically significant differences. In

total, then, there were 22 questions with statistically significant differences between teachers and administrators. The number of statistically significant differences between teachers and administrators was unexpected based on literature about these key stakeholder groups (Berliner, 2005a; Sarason, 1990; Cuban, 1998) as well as the perceptions of the teachers, administrators, and policymakers themselves (see Table 30, questions 4, 5, and 6). Teachers, administrators, and policymakers all expressed majority perceptions that teachers and policymakers held significantly different perspectives on public education; that administrators and policymakers held significantly different perspectives on public education; and that teachers and administrators "have the same perceptions about what is needed to improve public education" (EdRef_6). Yet collective results for individual survey questions, as shown by the ANOVA results, indicates that these perceptions are inaccurate. Closer examination of individuals questions by RSD teachers and administrators illustrates that these two groups strongly disagreed about many elements of RSD's current teacher evaluations, the ideals of a teacher evaluation system, education reform in general, the potential of SB 191, the passage of SB 191 and how to implement changes in general. These gaps in perception between teachers and administrators deserve closer examination.

The best means to explore teachers' and administrators' differences in perceptions, while also considering potential implications, was to combine quantitative and qualitative results. The pragmatist approach of mixed methods designs recognizes that quantitative data results must be taken as just one explanation of how teachers', administrators', and policymakers' perceive their roles in the implementation of SB 191

and not the sole explanation for these groups. Analysis of survey data embodies a postpositivist approach, for it assumes that answers can be found in the analysis of numbers. With a large enough sample size, the quantitative data may be useful to those charged with implementing the new teacher evaluation system in RSD, but it should never be taken as uninvestigated fact. The addition of qualitative data enriched this study by allowing teachers and administrators to articulate their perceptions for the "whys" behind the data. Interviews with practitioners with questions derived from the survey data, as well as participants' responses to the six open-ended survey questions, provided additional insights into practitioners' thinking. In this light, the best way to organize the quantitative and qualitative data was to use the codes generated during qualitative analysis—"teacher evaluations," "testing," "Senate Bill 191," and "teachers' voices in implementing school changes," with topics of "non-classroom educators" and "fear" incorporated throughout—and then review implications.

Teacher Evaluations. With respect to teacher evaluations, teachers and administrators disagreed on the merits of current teacher evaluations, what might constitute ideal teacher evaluations, and their understanding of the new evaluations. For RSD's existing teacher evaluation system, approximately 20% fewer teachers than administrators had faith in the evaluation's ability to encourage professional growth for teachers (C-TE_1a) and to assess teachers' overall teaching abilities (C-TE3a), even though 70% of teachers believed that their teaching had improved (C-TE_1) and agreed they had received meaningful feedback with these existing teacher evaluations (C-TE_2).

This gap between teachers' and administrator's perceptions of the effectiveness of the existing teacher evaluation may be explained by both Jake's (teacher) and Andrew's (administrator) individual reflections. Jake asserted that "[s]ome administrators definitely overestimate themselves. That's not up for debate." Andrew acknowledged that teachers "probably question" whether or not it was possible for administrators to conduct thorough, thoughtful evaluations under the existing system because of administrators' tendencies to observe a teacher for only one class, regardless of how many different classes a teacher may have. Additionally, a number of survey participants criticized the existing process for being superficial, calling it a mere "snapshot" or a "dog and pony show." One participant noted that "[v]iewing a teacher in a classroom for one hour even if its four times a year, cannot give an administrator an accurate reflection of a teacher's overall abilities." Other educators, particularly those serving students with special needs, shared their frustration with the evaluation process due to administrators' lack of knowledge—and, therefore, inability to provide meaningful feedback—about their area of specialty.

Even acknowledging these limitations, the fact that over two-thirds of teachers did believe that existing teacher evaluations were thorough and thoughtful was supported through responses to the open-answer survey questions. One participant wrote, "[i]f the administrator and teacher follow the evaluation process and steps, it works as intended and is a valuable tool." Another stated that his/her first evaluation in the district was "invaluable to me. I was amazed at how they broke down the craft into such detail."

In terms of ideal teacher evaluations, the survey data indicated that both teachers and administrators overwhelmingly agreed (94% and 97%, respectively) with the idea that teacher evaluations should "help teachers become better teachers" (IdealTE_2). But while 76% of administrators also believed that evaluations should be used to differentiate among teachers (IdealTE_1), fewer than 60% of teachers agreed with this idea. The reason for teachers' differing support for these two policies may best be summarized in interviewee Polly's (teacher) statement: "one is ideal and one is combative." Polly went on to explain that the concept of using teacher evaluations to help teachers grow was "just like in our classrooms: we take everybody from where they are and grow them… and there are definitely some best practice things that are learning, teachable, that can really change a teacher's success." In other words, the idea of using teacher evaluations to help teachers was consistent with teachers' approaches to their own classroom and, therefore, a comfortable concept. Additionally, Polly suggested that the idea of using teacher evaluations to differentiate among teachers was likely unpopular because it "goes back to the fear" factor of evaluations. Shelley (teacher) also articulated that the idea of helping teachers was positive and that evaluations "should be more of a coaching process," while the idea of using evaluations to sort educators into different groups felt more "punitive."

Faith in the ability of teacher evaluations, or any education reform, to improve student learning varied between teachers and administrators as well. Seventy percent of teachers and 80% of administrators believed that RSD's current evaluation system was linked to student learning (C-TE_2a). Approximately 93% or more of survey respondents

believed that teacher evaluations, in their ideal form, had the ability to improve student learning (IdealTE_3). However, 55% of teachers and 74% of administrators believed that education reforms are able to improve student learning (EdRef_1). This gap between teachers and administrators expanded when they were specifically asked about SB 191's reform efforts. In regards to the implementation of SB 191's new teacher evaluations in their school, only 15% of teachers and 32% of administrators agreed with the statement "Teachers at my school believe the new teacher evaluation system will be beneficial to students at my school" (SB191TE_3). It should be noted that 48% of administrators selected "unknown" as their answer choice to this question, illustrating a profound uncertainty about how teachers at their school felt. Similarly, just 36% of teachers, compared to 83% of administrators, agreed that the "reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way" (SB191_8).

Some participants voiced broad concerns about whether reforms can improve student learning in general. One of the more optimistic participant responses was "I think that reform CAN improve student learning if it is done well," but then he/she added, "I don't think that the current reforms are doing that." Another participant wrote many questions in connection to the idea of this policy improving student learning:

> State policymakers have no idea how difficult it will be to effectively evaluate every teacher, every year, at the high school level. Is it going to improve student learning to spend significantly more money on administrator salaries? Would this result in increased class size? Does the process just become meaningless because it just has to get done?

Practitioners appear ambivalent at best about student learning improving through the new teacher evaluations.

Testing. The quantitative data illustrated that RSD teachers and administrators also disagreed about how student assessment data should be incorporated into teacher evaluations. Sixty-six percent of teachers and 86% of administrators were comfortable with the linking of any form of assessment data to teacher evaluations (IdealTE_4). Seventy-nine percent of teachers and administrators believed an ideal teacher evaluation system would link student assessment data with teacher evaluations if the assessment was created by the teacher (IdealTE_6). However, only 31% of teachers agreed with using student data from state assessments (such as TCAP) while 72% of administrators supported it (IdealTE_5).

According to interviewees, these findings are logical. Teachers and administrators expected administrators to be more accepting of the inclusion of statewide assessments in teacher evaluations. Kate (teacher) acknowledged that administrators were more likely to be comfortable with statewide assessments because they are intended to be objective, an "equal measure that they can look at." She further explained that

> administrators feel, perhaps correctly, that teacher-made items vary a lot from teacher to teacher: what's hard in one class is not hard in another, the method of assignment varies a lot across the building and how those assessments are scored varies a lot across the building.

Shelley agreed with Kate, sharing that administrators can use statewide assessments "as evidence for a claim" about how teachers are performing. Jake noted that administrators were likely to be comfortable with statewide assessment data because they were used to it: "that's how administrators are evaluated already." Jake also noted that administrators do not see the comparison of teachers as "a personal thing," but instead as a way to be

"more objective" in their evaluations. Andrew noted many of the same things as Jake, stating that "as administrators, we're used to having data be part of our evaluation process."

As teacher interviewees reflected upon administrators' survey results, they also articulated several reasons behind teachers' negative responses to the use of statewide assessments. Kate succinctly asserted that "[t]he problem with statewide data, from a teacher's point of view, is that it is essentially mystery data." Kate explained that she used this description because the standards guiding the assessments were vague, teachers were not able to see the assessments, and the scoring system was unexplained. Polly went even further and asserted that teachers do not "trust" the statewide assessment; this lack of trust was highly problematic, she noted, for "if that test is part of our credentialing and evaluation, we need to trust the test as an excellent instrument." Jake articulated that teachers do not support the inclusion of student data from statewide assessments on teacher evaluations because students' lack of buy-in to perform well on such tests creates a "healthy amount of fear" that, in Jake's opinion, was "correctly placed." Andrew also noted students' lack of buy-in, but he approached the matter from a different angle.

> I think an administrator welcomes this idea [including student data on teacher evaluations] because it will help us with teachers having buy-in in students' performance on standardized tests because it's also going to affect them directly.

Participant responses to the open-ended survey questions revealed a plethora of comments regarding testing, particularly the use of statewide assessments, in connection to teacher evaluations. A few survey participants shared their positive support of the effort to link teacher evaluations to objective, statewide assessments. The vast majority

who wrote, though, railed against this. Some participants were concerned that the inclusion of student growth data from statewide assessments on teacher evaluations was a gross misappropriation of the data. Other participants, particularly those who work with students with special needs, expressed fear for their own jobs because their students were incapable of performing well on statewide assessments. Still other participants were concerned for pedagogical reasons, namely that the focus on student performance in a statewide assessment would shift the nature of learning in the classroom to content testing instead of critical thinking.

Senate Bill 191. The survey data indicated that few teachers or administrators were directly involved in passage of SB 191, yet teachers appeared to know less and expected less from the bill than administrators did. Forty-seven percent of teachers believed that they currently had an understanding of the legislation itself (SB191_1), 3% were involved with the passage of the legislation in some way (SB191_2), and 14 % supported the development of this bill (SB191_3) and were pleased with final version of SB 191 when it did pass (SB191_4). Importantly, a mere 13% of all teachers agreed with the statement "teachers at my school believe the new evaluation system will be beneficial to the teachers' professional practice at my school" (SB191TE_2). In other words, it is uncertain how well teachers understood the bill at the time of its passage or now but, given what they did know, they did not support it. Administrators expressed greater overall support for the bill. Despite only 17% of administrators noting that they had been involved with SB 191's legislative process (SB191_2), 50% stated that they had

supported the bill during its development (SB191_3) and 83% shared that they had a strong understanding of the legislation (SB191_1).

In terms of general familiarity with SB 191, the qualitative data reinforced the point that teachers had limited knowledge of the bill. Interviewee Polly noted that she has read SB 191 itself online, but that she knew of no peers who had done the same. Survey participants indicated that they remained unfamiliar with the bill by writing such things as, "[a]t this point I feel that most teachers are totally out of the loop when it comes to the new evaluation system" and "[w]e are still very confused!!!!" In fact, Anna (administrator) hypothesized that teachers' lack of knowledge about SB 191 was the probable cause of teachers' negativity toward the bill. Anna stated that

> people are running on their emotions right now. Yeah, absolutely. You're getting data from people that, I think, are emotionally thinking that, "I'm not very comfortable with this. I'm struggling with this."

She also suggested that, as educators learned more about the bill and the new teacher evaluations, they would become more supportive of these measures. Anna stated "once we get into the document [new teacher evaluation] and we really dig in deep and we help teachers understand how this is going to help them grow, you'll see a shift in your data."

According to the teachers interviewed, teachers' negativity regarding reforms included concerns about the perceived punitive nature of the reform efforts and the fact that elements of SB 191 identified needed areas of improvement in educational practices that did not align with teachers' own concerns about needed modifications. According to Jake, SB 191 was "about punishing teachers who don't get students to grow." Polly also

noted this perceived punitive nature, sharing her colleagues' concerns that the bill was more about catching teachers' errors than about an opportunity to grow.

> There's something scary about being watched all the time to see if we're slipping or doing something wrong or not updating our curriculum. I feel like I hear a lot of fear in teachers.

Kate's concern centered on the stated objectives of the bill compared to her understanding of the reforms to be enacted. Kate noted that

> the things identified in SB 191 don't pin-point what teachers think will really improve education or student learning…. So it's not necessarily that the [new teacher evaluation] standard is good or bad, it's just I think—from many teachers' point of view—it's not going to improve teaching.

Kate then provided two specific examples: the rubric and the standards. In regards to the new teacher evaluation rubrics themselves, Kate shared concerns about their lengthy nature: "the rubrics are slightly unwieldy, from a teacher's point of view, [and] they don't really get into what matters." Kate was also concerned that the rubrics were without substance sufficient to improve student learning. One standard that Kate found questionable was one that required all teachers to use the language of math in their classrooms. Kate wondered how, as an English teacher, her use of the word co-tangent was "going to influence how kids learn math, especially if I don't use that word correctly or I don't use it consistently…[I]t's not an English teacher talking about math that makes kids want to do math; it's what happens in math."

Related to these concerns, practitioners questioned the time demands of the new evaluations. Andrew articulated his reservations that SB 191's reforms would not improve student learning, in large part because the time necessary for administrators to

complete annual teacher evaluations for every teacher in the school will be extraordinarily demanding but not necessarily lead to changes in teachers' practices. "[W]hat I'm hearing is that it's going to be a huge amount of time." Numerous survey participants also shared their concern that the time demands of the new evaluations would completely consume administrators' time—and add marginally if at all to improvements in teachers' practices.

Teachers' voices in implementing school changes. Teachers and administrators agreed that neither group knew if or how teachers would be directly involved in the implementation of the new teacher evaluations and the accompanying changes, but that administrators would have direct involvement. Forty-six percent of teachers believed that they would play a role in implementing the new teacher evaluations (SB191TE_1), while 59% believed that teachers themselves were the central focus of SB 191's reform efforts (CF). Yet only 13% of teachers believed that teachers have primary responsibility for implementing the new reform, while 56% of teachers felt that administrators held the primary responsibility for enacting the reforms (PR). In contrast, 45% of administrators believed that teachers were the focus of reform efforts and 43% believed students were at the heart of the measure (CF). But 75% of administrators believed that administrators held the greatest responsibility for implementing the reforms (PR) and 89% of administrators agreed that they would play an active role in the implementation in their school (SB191TE_1).

There was ambiguity surrounding teachers' role in implementation, but teachers emphatically want to be included in the change process. Seventy-seven percent of

teachers believed that they generally respond well to changes at work (Change_2) but, more importantly, 98% of teachers responded affirmatively to the statement, "I respond well to change at work when I am involved with the creating the change" (Change_3). Teachers also responded strongly to the statement that there is "a right way and a wrong way to introduce new policies in a school" (Change_5). Administrators responded similarly, with responses averaging above 90% for all change questions. Administrators anticipated being directly involved with the changes coming to Colorado schools, though, and teachers apparently did not—despite their clear desire to be included.

Teachers' desire to be included in the change process was corroborated by the qualitative data. Interviewees articulated that the right way to conduct change was to include teachers in the process and the wrong way was for administrators to simply announce the changes. "I think the right way, and the way that we're trying to do it, is to talk about it as a team approach. We're going to figure it out together." Anna also shared that she was doing everything possible to make the process transparent, stating that she was doing what she could to make "teachers feel like they are a part of the process;" however, there is a difference between making one feel like he/she is part of a process and actually including him/her in the process. As one survey participant noted, "Engaging teachers as partners in the change is critical."

The interviewees also noted that school changes produced by fear were problematic on multiple levels. Polly asserted that "there's a lot of fear" in general in connection with the new teacher evaluations." Jake expressed that teachers were fearful

of consequences from the new teacher evaluations because many of the elements that will soon be impacting teacher evaluations are actually out of the teachers' control.

> I think in the abstract theory, it is the teachers who make the instructional decisions and they're the ones teaching in the classrooms every day, so they're the ones who need to be, whether you want to call it being held accountable or not, that's who this is directly geared towards… But when you get down to how it looks in each building and what matters in each building and which teachers are lauded for what they do and achieve, and which instructional models are chosen and which teachers are put in leadership positions and all of those types of decisions, those all come from the building administrator. And so the building administrator, by choosing people or programs or things that they highlight is setting the tone for what instruction needs to look like in the building.

Kate stated that teachers do not believe the bill has a chance to succeed "because they don't believe they have any influence over this process, really, and therefore they feel that improvement is not going to happen in schools." In other words, teachers believed that statewide mandated reforms will not make meaningful, lasting changes if teachers were not included in the change process for the reforms.

Implications

The guiding research question for this study was: How do Colorado teachers, administrators, and policymakers perceive their respective roles in the implementation of SB 191's new teacher evaluations? In brief, they hold differing perceptions. Teachers' perceived that they were the focus of SB 191's reform efforts, regardless of the bill's stated purpose to be improving student learning, and appeared resigned to the upcoming changes but were not embracing them. Administrators' perceived that the successful implementation of the new teacher evaluations fell to them naturally, "[b]ecause we're the school leaders," and were generally hopeful for the new teacher evaluations to

improve teachers' practice and students' learning. Policymakers were mostly hopeful about teacher evaluations being able to improve teachers' practices and student learning, but were cautious about asserting that SB 191's reforms would bring about such changes.

The data collected for this study and reviewed in this chapter appear consistent with the literature about why education reforms often fail to take root in schools. With two interviewees talking about teachers being jaded as well as numerous interviewees and survey participants articulating concerns about the anticipated high time demands of the new teacher evaluations, Tyack and Cuban's (1995) observation about overload negatively impacting reform efforts appears to be a relevant concern. The data also supported earlier literature stating that top-down reforms were particularly challenging for teachers when teachers were not given a legitimate voice in the process (Sarason, 1990; Hall & Hord, 1987), and teachers perceive that they do not have such a voice in the implementation of SB 191's new teacher evaluations at this point. Additionally, Sarason's (1990) argument for a shifting approach to education reforms, such that reforms consider both students and teachers, is consistent with the data, particularly in regards to those teachers who articulated doubt that the new evaluations would be able to make lasting improvements to student learning because teachers were not adequately consulted in the design and implementation phases.

From the critical social theory perspective, issues of power and voice remain relevant. Teachers and administrators are the practitioners primarily responsible for enacting the reforms resulting from SB 191's passage into law; 45% or more of everyone who completed the survey believed the central focus of the bill was teachers and 56% or

more agreed that the success of the bill will fall to administrators. Yet, it was policymakers who articulated what they believed needed to be improved in public education as well as how best to make this improvement; it was policymakers who had the power to create, pass, and enact SB 191. Teachers and administrators were marginally involved in the shaping of the bill and passage of the legislation. Administrators may anticipate and have a strong role to play in the implementation of the new teacher evaluations, but teachers' roles remain unclear, although the majority would like to be involved. This imbalance in power and voice between who sets the policies and who enacts the policies is problematic and, fundamentally, counterintuitive to creating successful, lasting school changes that improve students' learning across Colorado's public schools.

Given that data from this study corroborates both with the literature about challenges facing education reforms and with critical social theory's critiques of power and voice in public education reforms, it becomes important to explore both technical and adaptive options for those implementing SB 191's new teacher evaluations. Technical solutions are appropriate to consider in this case, for SB 191's new teacher evaluations are being implemented statewide starting in the 2013-14 school year; adaptive solutions, though, are preferred for exploring what long-term efforts may make school reforms more likely to develop lasting improvements.

Data from this study indicated that teachers did not fear higher-stakes evaluations because teachers disliked increased accountability as the media often portrays (Meyer, Bartels, & Fender, 2010; Meyer, 2010b), but that teachers were concerned about them for

three practical reasons. First, teachers were perplexed that SB 191 would create lengthier and more frequent teacher evaluations when one of the most consistent criticisms with the existing teacher evaluations was that they were superficial, "a dog and pony show," largely due to administrators' time constraints in conducting the evaluations. How, teachers wondered, could the new evaluation process provide more meaningful results—that might lead to improved pedagogy and, therefore, student learning—when administrators will have even less time to conduct the evaluations and provide feedback? Will the rigor of the new evaluations naturally erode because of the lack of time needed to conduct such evaluations? Second, teachers shared grave concerns over the inclusion of student learning growth data when precisely which assessments might be used to determine such student data and how the results will be processed remained unknown. Moreover, teachers questioned the feasibility of statewide assessments designed to assess student content knowledge to also be deemed valid and reliable means to assess teacher effectiveness. Third, RSD teachers appeared to be uncomfortable with the anticipated higher expectations for teachers without teachers having had a chance to participate in authentic discussions about how to implement the new teacher evaluations. Although administrators felt the implementation was "transparent" thus far, teachers' indicated that they largely felt excluded from the implementation process. Teachers desired to participate in creating the changes within their schools, but felt denied the opportunity to do so.

There are multiple technical solutions to address elements of these three concerns. To explore teachers' doubts related to having meaningful evaluations conducted within

the anticipated time constraints, it could be useful to begin by asking teachers: what would meaningful evaluations look like? A number of teachers commended existing peer evaluations for providing useful and timely recommendations for improved practice. One way to start the conversation about improving practice would be for administrators and/or district personnel to explore which elements of peer evaluations have proven particularly useful to teachers and find ways to include those elements more formally in the new teacher evaluations. Another might be to create a focus group of teachers from throughout the district to discuss recent findings (Gates Foundation, 2013) that student evaluations of teachers can be a positive addition to meaningful teacher evaluations. Yet a different way to address the concerns surrounding the anticipated time-demands of the new evaluation system would be for district leaders to confer with administrators about potential time constraints and discuss what district leaders might do to support these administrators during the first year or two of implementation of the new evaluations.

In terms of the fears associated with the use of student assessment data from statewide assessments, this may again fall to the district to explore how best to handle the incorporation of this additional information into teacher evaluations. Given that student learning growth data can only be compiled when there are at least two years of student assessment results, and Colorado's new assessments are designed to go into practice for the first time in the 2013-14 school year—the same year SB 191's new teacher evaluations begin statewide implementation—there will not be two years of student assessment data available, which is problematic. Rather than letting teachers' fears fester,

though, district personnel could confront this reality early, frequently, and directly, by explaining how the district intends to handle this anomaly.

Finally, to address teachers' expressed desires to participate in substantive conversations around the new evaluations' implementation efforts, it is incumbent upon school and district leadership to find authentic ways for more teachers to have real ownership of this reform. Perhaps something as straightforward as providing teachers with an opportunity to discuss their feelings regarding these changes in their own schools and with their own administrators would create positive results, for relationships are primary to successful changes (Wheatley, 1999) and honest conversations are capable of enhancing relationships. Given that teachers and administrators believed that they held similar perceptions about education reform matters, including teacher evaluations, but the data shows that they do not, it is all the more imperative for members of these two groups to engage in active conversations within their school communities.

Although these technical solutions might make inroads for certain teachers at certain schools, they would not address the larger, systemic issues underlying the new evaluations and the means to implementing reforms in general. Adaptive solutions are more complicated to consider because they often require paradigm shifts for stakeholders. However, they are also essential to consider if one hopes for lasting changes. One of Gramsci's and Bourdieu's shared critiques of public education was that it reproduced existing social stratification under the guise of equity and mobility. This critique appears prescient for SB 191. With SB 191, policymakers' power to set the direction and pace of education practices remains entrenched. Similarly, administrators' power over teachers

remains intact. In fact, this power actually expands under SB 191, for teacher evaluations will have higher-stakes than ever before, albeit under the stated intention of improving teachers' practices through the revised teacher evaluations and elimination of teacher tenure. In this study, teachers expressed the desire for increased involvement in enacting the changes in schools, but it is unclear what type of meaningful involvement they might be able to have under the current public education structure. In all the technical solutions just presented, it is district personnel and administrators who would provide opportunities to teachers for dialogue; there are no conversations among equals. Moreover, the ability for anything productive to come out of the recommended conversations would rest, unequally, with district personnel and administrators. If Sarason's proclamation that teachers need to be considered and grown as much as students for public education reforms to make lasting, positive changes, then it remains unlikely that SB 191 will be able to generate such changes for teachers' power has been reduced, rather than strengthened, by this legislation.

Adaptive solutions might entail questioning structures as large as public education in general, asking such questions as: Does the hierarchical approach within schools, such as administrators having power over teachers, ultimately serve schools and students well in 2013? Do the skills students need in the 21st century match the structure of schools carried over from the 20th century? These larger, meta-questions have no simple solutions, but asking them at least opens the door to "a universe of possibility" that otherwise remains closed (Zander and Zander, 2000, p. 15).

Areas for Future Research

This study provides multiple avenues for future studies.

One of the reasons SB 191 has garnered national attention is its formal elimination of teacher tenure. Survey participants mentioned job stability and tenure frequently in the open-ended response questions, but there were no survey questions specific to this topic. An exploration of perceptions directly related to teacher evaluations and to teacher tenure, or the lack thereof, could be a direction for further research.

Survey data indicated that the common perception of administrators and policymakers having dissimilar views education reform may be inaccurate, for the responses of administrators and policymakers to survey questions were often similar. However, the low response rate for policymakers makes it difficult to know if these findings are accurate or anomalous. Additional research into policymakers' perspectives on education reform could be illuminating.

Many survey participants articulated concerns related to the inclusion of student learning growth data in teacher evaluations, particularly when the data came from statewide assessments. Some raised the issue of statewide assessments being normed to assess both student knowledge and teacher effectiveness; a study exploring Colorado's new statewide assessments' validity and reliability in connection to determining teacher effectiveness could be useful. From a different angle, it could also be valuable to explore if and how district-wide teacher evaluation results (i.e., the percentage of teachers with effective versus ineffective ratings) change as a result of the inclusion of student learning growth data in teacher evaluations.

This study did not explicitly explore the role of district personnel, including superintendent and school boards, in the implementation of SB 191's reform efforts. Yet both teachers and administrators indicated that the implementation of new teacher evaluations was being coordinated at the district level. A study exploring the perceptions of district personnel and school boards in regard to teacher evaluations, education reform, and implementing reforms could be helpful in understanding the full complexity of these matters.

One teacher interviewee noted that this study's survey encouraged teachers and administrators to ask "what do we believe *will* make us better?" (italics reflecting the author's emphasis). As this study's literature review noted, few studies have solicited teachers' and administrators' perceptions; from both a district and a statewide policy perspective, it could be eye-opening to see the ideas teachers and administrators would generate if a study were conducted with this as the guiding research question.

Concluding Ideas

Teachers, administrators, and policymakers are three of the central stakeholders in public education throughout the United States. Each group holds a different but critical role and these roles are interconnected. Yet, members of the three groups rarely communicate their perceptions to the others. This must shift if meaningful reforms are going to be legislated and implemented with fidelity.

As the data from this study illustrated, the perceptions of teachers, administrators, and policymakers often differ from how each group sees itself or how media portrays the group. Everyone involved expressed the desire to improve student learning through

teacher evaluations, but the means to do so remain ambiguous. Across these three stakeholder groups, belief in SB 191's new teacher evaluations to create such positive changes for student learning is reserved, at best. It is clear, though, that those with the greatest experience in schools and insights into educational practices currently have less voice and power than those without. As long as this power imbalance remains, it is doubtful that substantive changes will occur through this reform. Finding a way to bring these groups together on equal footing at school, district, and state levels, though, could yet prove powerful by forging trust and setting a lasting path forward toward improving student learning.

References

Alcias, E. R. (2005). Toward an objective evaluation of teacher performance: The use of variance partitioning analysis, VPA. *Educational Policy Analysis Archives, 13.*

Amrein-Beadersley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*, 65-75. doi:10.3102?0013189X08316420

Auge, K. (2012, November 1, 2012). Failing grades cut off grants. *The Denver Post*, pp. 1-10A.

Barnum, M. (2012, August 26). Blueprints for teacher evaulations. *The Denver Post*, p. 4D.

Berliner, D. C. (2005a). Our impoverished view of educational reform. *Teachers College Record.*

Berliner, D. C. (2005b). The near impossibility of testing for teacher quality. *Journal of Teacher Education, 56*(3), 205-213. doi:10.1177/0022487105275904

Bernstein, H. (1936, May 24). Security of the teacher in his job. *New York Times*

Berry, B. (2010). Getting "real" about teaching effectiveness and teacher retention. *Journal of Curriculum and Instruction, 4*(1), 1.

Berry, B. (2011). *Teaching 2030: What we must do for our students and our public schools--now and in the future*. Teachers College Press. www.tcpress.com

Bobko, P. (2001). *Correlation and regression: Applications for industrial organizational psychology management* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Bohman, J. (2005). Critical theory In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2012 ed.)

Bolman, L. G., & Deal, T. E. (2008). *Reframing organizations: Aristry, choice, and leadership*. San Francisco, CA: Jossey-Bass.

Bonk, W. J. (2010). The Colorado growth model: Using norm- and criterion- referenced growth calculations to ensure that all students are held to high academic standards. *Colorado Department of Education*. Retrieved on October 12, 2012, from http://www.cde.state.co.us/Accountability/Downloads/GrowthStandardsAccountability.pdf

Briggs, D., & Domingue, B. (2011). *Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles unified school district teachers by the Los Angeles times.* Boulder, CO: National Education Policy Center.

Brill, S. (2011). *Class warfare: Inside the fight to fix America's schools.* New York: Simon and Schuster.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood.* (No. 17699). Cambridge, MA: National Bureau of Economic Research.

Colorado Department of Education, Approved educator preparation programs. Retrieved November 27, 2012, from http://highered.colorado.gov/Academics/TeacherEd/forms/ApprovedProgramsX.pdf

Colorado Department of Education, Educator Effectiveness. Retrieved December 11, 2012, from http://www.cde.state.co.us/EducatorEffectiveness/FAQs.asp#Other_Licensed_Professionals

Colorado Department of Education, Educator Effectiveness. (2012). Rubric for evaluating Colorado's teachers. Retrieved April 19, 2013, from http://www.cde.state.co.us/EducatorEffectiveness/downloads/SMES%20-%20Teacher/RubricforEvaluatingCOTeacherFINAL.pdf

Couto, R. A. (2003). Review essay: Community-based research: Celebration and concern. *Michigan Journal of Community Service Learning, Summer*, 69-74.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage Publishers, Inc.

Cuban, L. (1998). How schools change reforms: Redefining reform success and failure. *Teachers College Record, 99*(3), 453-477.

Danzer, G. A. (2003). *The Americans*. Boston: McDougal Littell, Inc.

Dewey, J. (1916). *Democracy and education*. New York: Macmillian Company.

Dillon, S. (2010, December 7). Top test scores from Shanghai stun educators. *New York Times*

Duhigg, C. (2012). *The power of habit: Why we do what we do in life and business*. New York, NY: Random House.

Education Week. (2013, January 10). State and national grades issued for education performance, policy; U.S. earns a C-plus, Maryland ranks first for fifth straight year. *Education Week*

Engdahl, T. (2011). Districts eager for SB 191 pilot. *EdNews Colorado.* www.ednewscolorado.org

Forte, E. (2010). Examining the assumptions underlying the NCLB federal accountability policy on school improvement. *Educational Psychologist, 45*(2), 76-88.

Fowler, F. J. (1995). In Susan McElroy (Ed.), *Improving survey questions: Design and evaluation*. Thousand Oaks, CA: Sage Publications, Inc.

Fowler, F. J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Friedman, H. H., & Amoo, T. (1999). Rating the ratings Scale**.** *Journal of Marketing Management, 9*(3), 114-123.

Friedman, T. L. (2011, November 19). How about better parents? *New York Times*

Fullan, M. (2006). *Turnaround leadership*. San Francisco, CA: Jossey-Bass.

Gates Foundation. (2013). *Feedback for better teaching: Nine principles for using measures of effective teaching*. Seattle, WA: Bill and Melinda Gates Foundation. Retrieved on May 17, 2013, from: http://www.metproject.org/downloads/MET_Feedback%20for%20Better%20Teaching_Principles%20Paper.pdf

Glazerman, S., Loeb, S., Goldhaber, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The role of value-added.* Washington, D. C.: Brown Center on Education Policy at Brookings.

Goodnough, A. (2011, February 25). Mayor tries to reassure providence teachers as furor grows over firing notices. *New York Times*

Graham, P. A. (2005). *Schooling America: How the public schools meet the nation's changing needs*. New York: Oxford University Press.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-methods evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Haley, D. (2010). Democrats gone wild. *The Denver Post*, p. D-03.

Hall, G. E., & Hord, S., M. (1987). *Change in schools: Facilitating the process*. Albany, NY: State University of New York Press.

Hargreaves, A., & Fink, D. (2006). *Sustainable leadership*. San Francisco, CA: Jossey-Bass.

Hargreaves, A., & Shirley, D. (2008). The fourth way. *Educational Leadership*, 56-61.

Hatlestad, L. (2012). Forced resignation. *5280: The Denver Magazine, 20*(2), 110.

Hlebowitsh, P., & Tellez, K. (1997). *American education: Purpose and promise*. New York: West/Wadsworth.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly, 24*(4), 602-611.

Johnson, S. M. (2009). *How best to add value? strike a balance between the individual and the organization in school reform.* (No. 249). Washington, D.C.: Economic Policy Institute.

Johnston, M. Online newsletter from Senator Mike Johnston. Retrieved November 16, 2010, from http://www.stand.org/Document.Doc?id=2650

Johnston, M., & Barker, L. (2011, September 16). *University of Denver roundtable discussion: Senate bill 10-191*

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains.* Bill and Melinda Gates Foundation: Bill and Melinda Gates Foundation.

Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly, 45*(1), 34-70. doi:10.1177/0013161X08327549

Kristof, N. (2012). Students over unions. *New York Times*

Labaree, D. F. (2010). *Someone has to fail: The zero-sum game of public schooling*. Cambridge, MA: Harvard University Press.

Lamm, R., Romer, R., Owens, B., & Ritter, B. J. (2010). Education bill's great promise. *The Denver Post*, pp. B-11.

Levinson, B. A. U., Gross, J. P. K., Hanks, C., Dadds, J. H., Kumasi, K. D., Link, J., & Metro-Roland, D. (2011). *Beyond critique: Exploring critical social theories and education*. Boulder, Colorado: Paradigm Publishers.

Mann, H. (1872). *Annual reports on education*. Boston, MA: Lee and Shepard.

Mariano, L. T., McCaffrey, D. F., & Lockwood, J. R. (2010). A model for teacher effects from longitudinal data without assuming vertical scaling. *Journal of Educational and Behavioral Statistics, 35*, 253. doi:10.3102/1076998609346967

Martineau, J. (2006). Distorting value added: The use of longitudinal, vertically scaled student achievement data for growth-based, value-added accountability. *Journal of Educational and Behavioral Statistics, 31*, 35-62. doi:10.3102/10769986031001035

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*, 67-101. doi:10.3102/10769986029001067

McGrane, M. (2009). *Colorado's growth model.* Unpublished manuscript.

McLaren, P., Martin, G., Farahmandpur, R., & Jaramillo, N. (2004). Teaching in and against the empire: Critical pedagogy as revolutionary praxis. *Teacher Education Quarterly, 31*(1), 131-153.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Meyer, J. P. (2010a, February 19). DPS teacher evaluations don't mirror quality, group says. *The Denver Post*

Meyer, J. P. (2010b, May 13). “Positive model" for states weighing teacher reforms. *The Denver Post*, p. A-01.

Meyer, J. P. (2010c). Teacher bill sees division - A house committee debates senate legislation that would alter how tenure is obtained and kept. *The Denver Post*, p. A-01.

Meyer, J. P., Bartels, L., & Fender, J. (2010, May 12). Legislature 2010: Teacher-evaluation reform, bill nears passage, dems filibuster; others cross over to set up final vote. *The Denver Post*, p. A-01.

New Millennium Teachers. (2012). *Making teacher evaluation work for students: Voices from the classroom.* New Millennium: An Initiative of the Center for Teaching Quality.

Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives, 14*(1).

Noddings, N. (2001). Care and coercion in school reform. *Journal of Educational Change, 2*(1), 35-43.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education, 33*(3), 301-314. doi:10.1080/02602930701293231

OECD. (2011). *Strong Lessons from PISA for the united states, strong performers and successful reformers in education.* OECD Publishing.

Ooms, A. (2011, July 25). A tipping point for democrats on education. *The Denver Post*, p. A-19.

Paul, A. M. (2012). It’s not me, it’s you. *New York Times*

Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers.* Boston, MA: National Board on Educational Testing and Public Policy.

Pena, F., & Zeller, L. H. (2010, May 7). Pass the teacher effectiveness bill. *The Denver Post*, p. B-11.

Phillips, D. C. (2004). Two decades after: "after the wake—postpositivistic educational thought." *Science & Education, 13*(1-2), 67-84.

Ravitch, D. (2010). *Death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Reason, P., & Bradbury, H. (2006). *Handbook of action research*. Los Angeles, CA: Sage Publications.

Resmovits, J. (2012, December 11). International tests show East Asian students outperform world as U.S. holds steady. *The Huffington Post*

Rich, M. (2012, 9/11/12). National schools debate is on display in Chicago. *The New York Times*

Robinson, S. K. (2010). TED: Changing education paradigms. Retrieved November 13, 2012, from http://www.ted.com/talks/ken_robinson_changing_education_paradigms.html

Robles, F. (2007). *Veteran teachers' perspectives on teacher evaluation and how they want to be evaluated.* (Unpublished PhD). University of the Pacific, Stockton, CA.

Robles, Y. (2011). Schools question teacher ratings - local officials believe state's proposed rules will limit districts' options for evaluation. *The Denver Post*, p. B-04.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247-252.

Sahlberg, P. (2010). *The secret to Finland's success: Educating teachers*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Sarason, S. (1990). *The predictable failure of education reform: Can we change course before it's too late?*. San Francisco, CA: Jossey-Bass Publishers.

Sass, T. R. (2008). The stability of value-added measures of teacher quality and implications for teacher compensation policy. *National Center for Analysis of Longitudinal Data in Education Research, (4)*, 1-7.

Senate Bill 10-191: Great teachers and leaders act, 191, (2010). 1: 22-9-102. http://www.cde.state.co.us/cdedepcom/download/pdf/SB10-191.pdf

Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that learn: A fifth discipline fieldbook for educators, parents, and everyone who cares about education*. New York: Doubleday.

Sergiovanni, T. J., & Starratt, R. J. (2002). *Supervision: A redefinition* (7th ed.). Boston, MA: McGraw-Hill Higher Education.

Spring, J. (1997). *The American school: 1642-1996* (4th ed.). New York: McGraw-Hill.

Stone, C. A., & Lane, S. (2003). Consequences of a state accountability program: Examining relationships between school performance gains and teacher, student, and school variables. *Applied Measurement in Education, 16*(1), 1-26.

Strand, K., Marullo, S., Cutforth, N., Stoecker, R., & Donohue, P. (2003). Principles of best practice for community-based research. *Michigan Journal of Community Service Learning, Spring*, 5-15.

Tashakkori, A., & Teddlie, C. (2007). Introduction to mixed method and mixed model studies in the social and behavioral sciences. In V. Plano Clark, & J. W. Creswell (Eds.), *The mixed methods reader* (p. 7). Thousand Oaks, CA: Sage Publications, Inc.

Torres, M. S. (2004). Best interests of students left behind? Exploring the ethical and legal dimensions of United States federal involvement in public school improvement. *Journal of Educational Administration, 42*(2), 249-269.

Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: Reflections on a century of public school reform*. Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1993). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*(3), 453-479.

U. S. Department of Education. (2009). *Race to the top program executive summary*.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2011). *Final report on the evaluation of the growth model pilot project.* Washington, D.C.: U.S. Department of Education.

University of Wisconsin-Madison. Qualtrics survey hosting service, survey best practices. Retrieved January 5, 2013, from http://survey.wisc.edu/BestPractices.html

Web Center for Social Research Methods. Research methods knowledge base, the T-test. Retrieved January 6, 2013, from http://www.socialresearchmethods.net/kb/stat_t.php

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York City: The New Teacher Project.

Wheatley, M. (1999). *Leadership and the new science: Discovering order in a chaotic world* (2nd ed.). San Francisco: Berrett-Koehler Publishers.

White House Press Release. (November 4, 2009). Fact sheet: Race to the top. Retrieved August 21, 2012, from http://www.whitehouse.gov/the-press-office/fact-sheet-race-top

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom content effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57-67.

Zander, R. S., & Zander, B. (2000). *The art of possibility: Transforming professional and personal life.* New York, NY: Penguin Books.

Zezima, K. (2010, February 24). A vote to fire all teachers at a failing high school. *New York Times*

Appendices

Appendix A

Perceptions about the Implementation of SB 191's New Teacher Evaluations

Survey

Thank you in advance for taking the time to complete this survey as part of my dissertation research with the University of Denver. I am studying the implementation of Colorado Senate Bill 10-191: The Great Teachers and Leaders Act. It was proposed as legislation in April, 2010 and became law in May, 2010. The new teacher evaluation system is one of the many changes resulting from the passage of this bill.

I am eager to understand how three key stakeholders in education reform—teachers, administrators, and policymakers—each perceive their role in the implementation of Colorado Senate Bill 10-191's new teacher evaluation system. Your participation is critical to helping give voice to the perspectives of your particular group.

Based on timed pilot testing, it should take approximately 10 minutes to complete this survey. Participation in this study is voluntary. By clicking to the following page and continuing with this study, you indicate informed consent to participate in this study.

If you would like additional information about this study and/or your rights as a participant, please email: sbridich@du.edu. This survey was approved by the University of Denver's Institutional Review Board for the Protection of Human Subjects in Research on 10/04/2012.

Survey results are anonymous and confidential. If you would like to receive a summary of the findings from the survey, please email: sbridich@du.edu and results will be sent to you following the study's completion.

Again, thank you for your participation.

Sarah Melvoin Bridich

*** page split ***

Q17 Please check your current position.

- ❍ Teacher (1)
- ❍ Administrator (2)
- ❍ Policymaker (3)

Answer If Please check your current position. Teacher Is Selected

Q42 Please select if you are a:

- ❍ Classroom teacher in a TCAP tested subject (1)
- ❍ Classroom teacher in a non-TCAP tested subject (2)
- ❍ Other educator in school/licensed professional (3)

Answer If Please select if you are a: Other educator in school Is Selected

Q28 Please type your current position below:

Answer If Please check your current position. Administrator Is Selected

Q27 Please select if you are a:

- ❍ Building Administrator (1)
- ❍ District Administrator (2)

*** page split ***

Answer If Please check your current position. Teacher Is Selected

Q24 Instructions: Please answer the following questions based on your personal experiences with the current teacher evaluation process at your school. The current teacher evaluation system in my school:

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)
has improved my teaching practices. (1)	❍	❍	❍	❍
has provided me with meaningful feedback. (2)	❍	❍	❍	❍
has made me a better teacher overall. (3)	❍	❍	❍	❍

*** page split ***

Answer If Please check your current position. Policymaker Is Not Selected

Q1 Instructions: Please answer the following questions based on your personal experiences with the current teacher evaluation process at your school. The current teacher evaluation process in my school:

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)
encourages professional growth for teachers. (1)	❍	❍	❍	❍
is linked to student learning. (2)	❍	❍	❍	❍
is able to assess teachers' overall teaching abilities. (3)	❍	❍	❍	❍
is a high stress process for teachers. (4)	❍	❍	❍	❍
is a high stress process for administrators. (5)	❍	❍	❍	❍
allows for thorough reviews of teachers' overall teaching abilities. (6)	❍	❍	❍	❍
allows for thoughtful reviews of teachers' overall teaching abilities. (7)	❍	❍	❍	❍

Answer If Please check your current position. Policymaker Is Not Selected

Q35 If you have additional information and/or comments to share about this topic, please use the space below.

*** page split ***

Q2 Instructions: Please answer the following questions based on how you believe the teacher evaluation process should ideally function. Teacher evaluations are best when they:

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)
distinguish strong teachers from weak teachers. (1)	❍	❍	❍	❍
help teachers become better teachers. (2)	❍	❍	❍	❍
have the ability to improve student learning. (3)	❍	❍	❍	❍
are linked to student assessment data of any form. (4)	❍	❍	❍	❍
are linked to student assessment data created by the state department of education (i.e., CSAPs). (5)	❍	❍	❍	❍
are linked to student assessment data created by the teacher. (6)	❍	❍	❍	❍

Q37 If you have additional information and/or comments to share about this topic, please use the space below.

*** page split ***

Q3 Instructions: Please answer the following questions based on your opinion.

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)
I view change as an opportunity for growth. (1)	❍	❍	❍	❍
I respond well when policies change at work. (2)	❍	❍	❍	❍
I respond well to change at work when I am involved with creating the change. (3)	❍	❍	❍	❍
I respond well to change at work when informed of what new policies will be and I must learn the new policies. (4)	❍	❍	❍	❍
There is a right way and a wrong way to introduce new policies in a school. (5)	❍	❍	❍	❍

Q38 If you have additional information and/or comments to share about this topic, please use the space below.

*** page split ***

Q4 Instructions: Please answer the following questions based on your understanding of educational reforms in general.

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)
Education reforms improve student learning. (1)	❍	❍	❍	❍
Education reforms view teachers as knowledgeable professionals. (2)	❍	❍	❍	❍
Education reforms view building administrators as knowledgeable professionals. (3)	❍	❍	❍	❍
Policymakers (i.e., legislators, lobbyists) and teachers have the same perceptions about what is needed to improve public education. (4)	❍	❍	❍	❍
Policymakers (i.e., legislators, lobbyists) and building administrators have the same perceptions about what is needed to improve public education. (5)	❍	❍	❍	❍

Teachers and building administrators have the same perceptions about what is needed to improve public education. (6)	❍	❍	❍	❍

Q5 Teachers should lead the implementation of reforms at the: (check all that apply)
- ❑ Building Level (1)
- ❑ District Level (2)
- ❑ State Level (3)
- ❑ None of the above (4)

Q7 Building-administrators should lead the implementation of reforms at the: (check all that apply)
- ❑ Building Level (1)
- ❑ District Level (2)
- ❑ State Level (3)
- ❑ None of the above (4)

Q8 Policymakers should lead the implementation of reforms at the: (check all that apply)
- ❑ Building Level (1)
- ❑ District Level (2)
- ❑ State Level (3)
- ❑ None of the above (4)

Q39 If you have additional information and/or comments to share about this topic, please use the space below.

*** page split ***

Answer If Please check your current position. Policymaker Is Not Selected

Q9 Instructions: Please answer the following questions based on your understanding of SB 191's new teacher evaluation system.

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)	Don't Know (5)
I will play an active role in the implementation of my school's new teacher evaluation system. (1)	❍	❍	❍	❍	❍
Teachers at my school believe the new teacher evaluation system will be beneficial to the teachers' professional practice at my school. (2)	❍	❍	❍	❍	❍
Teachers at my school believe the new teacher evaluation system will be beneficial to students at my school. (3)	❍	❍	❍	❍	❍

*** page split ***

Q10 Instructions: Please answer the following questions based on your understanding of SB 191's new teacher evaluation system.What roles would be most helpful for teachers to play during the implementation of SB 191 at their school? (Check all that apply)

- ❑ Design review processes and timelines (1)
- ❑ Provide input for evaluation criteria (2)
- ❑ Provide support to other teachers (3)
- ❑ Develop common student assessments (4)
- ❑ Other (5)

Q13 Which group has primary responsibility for the successful implementation of the new teacher evaluation system? (select only one)

- ❍ Teachers (1)
- ❍ Building Administrators (2)
- ❍ District Personnel (3)
- ❍ Policymakers (i.e., legislators, lobbyists, etc.) (4)
- ❍ Other (5)
- ❍ Unsure (6)

Q40 If you have additional information and/or comments to share about this topic, please use the space below.

*** page split ***

Q14 Instructions: Please answer the following questions based on your personal experiences with Colorado Senate Bill 10-191.

	Completely Disagree (1)	Generally Disagree (2)	Generally Agree (3)	Completely Agree (4)	Not Applicable (5)
I have a strong understanding of the legislation SB 191. (1)	❍	❍	❍	❍	❍
I was highly involved in the development of SB 191 during its implementation phase. (2)	❍	❍	❍	❍	❍
I supported SB 191 during its development. (3)	❍	❍	❍	❍	❍
I was pleased with the final version of SB 191 as it was passed. (4)	❍	❍	❍	❍	❍
The media portrayed teachers fairly during the bill's passage. (5)	❍	❍	❍	❍	❍
The media portrayed administrators fairly	❍	❍	❍	❍	❍

during the bill's passage. (6)					
The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage. (7)	❍	❍	❍	❍	❍
The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way. (8)	❍	❍	❍	❍	❍

Q15 The central focus of SB 191 is

- ❍ Students (1)
- ❍ Teachers (2)
- ❍ Building Administrators (3)
- ❍ Policymakers (4)

Q16 Check all sources from where you obtained your knowledge of SB 191:

- ❑ CEA representatives or materiasl (mailing, email, etc.) (1)
- ❑ AFT representatives or materiasl (mailing, email, etc.) (2)
- ❑ Your school district's administrators (3)
- ❑ Your building administrators (4)
- ❑ Professional development workshop (5)
- ❑ Faculty meeting agenda item (6)
- ❑ Fellow teachers (7)
- ❑ Fellow policymakers (8)
- ❑ Media: radio, television, print (9)
- ❑ Internet Information (10)
- ❑ Other (11)

Q41 If you have additional information and/or comments to share about this topic, please use the space below.

**** page split ****

Q29 Instructions: For the following questions, please answer based on your personal experience.Your gender:

- ❍ Male (1)
- ❍ Female (2)

Q23 Your ethnicity:

- ❑ Asian (1)
- ❑ African-American/Black (2)
- ❑ Hispanic/Latino (3)
- ❑ Native American or Alaskan Native (4)
- ❑ Native Hawaiian or Pacific Islander (5)
- ❑ White (6)

Answer If Please check your current position. Policymaker Is Not Selected

Q19 Current School Type:

- ❍ Elementary School (1)
- ❍ K-8 (2)
- ❍ Middle School (3)
- ❍ High School (4)
- ❍ Other (5)

Answer If Please check your current position. Teacher Is Selected

Q26 Please select your license specialty. If you have more than one, please select just one.

- ❍ Art (1)
- ❍ Business and Marketing Education (2)
- ❍ Counselor (3)
- ❍ Drama (4)
- ❍ Early Childhood Education (5)
- ❍ Elementary Education (6)
- ❍ English Language Arts (7)
- ❍ Family and Consumer Studies (8)
- ❍ Foreign Language (9)
- ❍ Gifted and Talented Specialist (10)
- ❍ Health (11)
- ❍ Instructional Technology (12)
- ❍ Librarian (13)
- ❍ Culturally & linguistically diverse education (14)
- ❍ Culturally & linguistically diverse bilingual education (15)
- ❍ Marketing education (16)
- ❍ Mathematics (17)
- ❍ Music (18)
- ❍ Nurse (19)
- ❍ Occupational Therapist (20)
- ❍ Orientation and Mobility Specialist (21)
- ❍ Physical Education (22)
- ❍ Psychologist (23)
- ❍ Reading (24)
- ❍ Science (25)
- ❍ Social Studies (26)
- ❍ Social Worker (27)
- ❍ Special Education (28)
- ❍ Speech (29)
- ❍ Technology Education (30)
- ❍ Trade and Industry Education (31)
- ❍ Agriculture and Renewable Natural Resources (32)
- ❍ Audiologist (33)
- ❍ Physical Therapist (34)

- ❍ Teacher Librarian (35)

Answer If Please check your current position. Policymaker Is Not Selected

Q43 Please select the type of license you currently have.
- ❍ Initial (1)
- ❍ Professional (2)
- ❍ Alternative (3)

Q20 Total years in education
- ❍ 1 (1)
- ❍ 2 (2)
- ❍ 3 (3)
- ❍ 4-6 (4)
- ❍ 7-15 (5)
- ❍ 16-25 (6)
- ❍ 26+ (7)
- ❍ None of the above (8)

Answer If Please check your current position. Policymaker Is Not Selected

Q22 Total years in district
- ❍ 1 (1)
- ❍ 2 (2)
- ❍ 3 (3)
- ❍ 4-6 (4)
- ❍ 7-15 (5)
- ❍ 16-25 (6)
- ❍ 26+ (7)

Answer If Please check your current position. Policymaker Is Selected

Q30 Total years in elected office (all positions combined)

- ❍ 1 (1)
- ❍ 2 (2)
- ❍ 3 (3)
- ❍ 4-6 (4)
- ❍ 7-15 (5)
- ❍ 16-25 (6)
- ❍ 26+ (7)

Answer If Please check your current position. Policymaker Is Not Selected

Q31 Approximately what percentage of students in your school are students of color (e.g., African American, Latino/a, American Indian, Asian, Pacific Islander)?

- ❍ 0%-20% (1)
- ❍ 21%-40% (2)
- ❍ 41%-60% (3)
- ❍ 61%-80% (4)
- ❍ 81%-100% (5)
- ❍ Unsure (6)

Answer If Please check your current position. Policymaker Is Not Selected

Q32 Approximately what percentage of students in your school speak a language other than English as their primary language?

- ❍ 0%-20% (1)
- ❍ 21%-40% (2)
- ❍ 41%-60% (3)
- ❍ 61%-80% (4)
- ❍ 81%-100% (5)
- ❍ Unsure (6)

Answer If Please check your current position. Policymaker Is Not Selected

Q33 Approximately what percentage of students in your school have disabilities?

- ❍ 0%-20% (1)
- ❍ 21%-40% (2)
- ❍ 41%-60% (3)
- ❍ 61%-80% (4)
- ❍ 81%-100% (5)
- ❍ Unsure (6)

Answer If Please check your current position. Policymaker Is Not Selected

Q44 Approximately what percentage of students in your school are on Free and Reduced Lunch (FRL)?

- ❍ 0%-20% (1)
- ❍ 21%-40% (2)
- ❍ 41%-60% (3)
- ❍ 61%-80% (4)
- ❍ 81%-100% (5)
- ❍ Unsure (6)

Appendix B

Guiding Areas for Survey Creation

These ideas shaped questions for the survey but were not shared with participants.

Area	Explanation
Area 1: Perceptions of the current teacher evaluation system	All teachers and administrators have personal experiences tied to their understanding of the current teacher evaluation system. Before exploring perceptions of SB 191, it is important to get a sense of how the current/previous teacher evaluation is/was perceived.
Area 2: Perceptions of the teacher evaluation process, in general	It is important to understand how individuals perceive the role of teacher evaluations in general within the scope of public education.
Area 3: Perceptions of organizational change	Change is challenging for most people and organizations. How individuals respond to change can be indicative of how individuals will incorporate or reject reform practices. Therefore, asking general questions about change may correlate both to participants' responses to specifics about SB 191 and to a trend larger than how individuals/groups feel about SB 191's new evaluation system.
Area 4: Perceptions of education reforms, in general	Understanding how groups of stakeholders perceive education reforms in general could be helpful in implementing reforms.
Area 5: Perceptions of the implementation of SB 191's new teacher evaluation system	Who is involved in the implementation of any education reform impacts the chance of successful implementation. Understanding how each of these three key stakeholders perceives their role, and the role of the others, could be helpful in implementing this reform.
Area 6: Perceptions of SB 191, in general	Colorado Senate Bill 10-191 received much media attention, particularly in the heated days before the bill passed. Although the bill's passage was in May 2010, precisely what Colorado's educators know about it is unknown. Because the passage of the bill was so heated, it would be useful to know how participants understand the bill itself and what influenced this understanding.

Appendix C

Pilot Testing Questions

The following were the questions asked during the pilot testing of the survey.

General:

1. Are there any questions which made you want to stop taking the survey?
2. Were you confused by your answer choices? Did “Generally Disagree” make sense, or would you have preferred “Somewhat Disagree”?

Clarity:

3. Did you have content questions?
4. Did you have vocabulary questions?
5. Does anything need clarification in general?

Content:

6. Did any questions surprise?
7. Did any questions seem repetitive?

Appendix D

Pilot Testing Consent Form

I am asking you to participate in the piloting of a research study. This form is designed to give you information about this study. I will describe this study to you and answer any of your questions.

Project Title: *Implementation Roles: Teachers', Administrators, and Education Policymakers' Perceptions of their Roles in the Implementation of Colorado Senate Bill 10-191's New Teacher Evaluation System*

Principal Investigator: Sarah M. Bridich
University of Denver, Morgridge College of Education, Education Administration and Leadership
sbridich@du.edu

Faculty Advisor: Dr. Lyndsay Agans
University of Denver, Morgridge College of Education, Education Administration and Leadership
Lagans@du.edu

What the study is about
The purpose of this research is to explore Colorado's teachers', administrators' and policymakers' perceptions of their role in the implementation of Colorado Senate Bill 10-191, specifically the new teacher evaluation system.

What I will ask you to do
I will ask you to complete the draft survey and then to share your impressions of the survey.

Risks and discomforts
I do not anticipate any physical, legal, or economic risks associated with this study. The conversation will be about implementation reform, particularly the new teacher evaluation system. That said, teacher evaluation can be an emotional topic and, consequently, there could be emotional risks associated with participation. Every effort will be made by the moderator (myself) to make participants as comfortable as possible discussing the content. Additionally, no one will be asked to share personal teacher evaluation stories. Finally, participants could feel a social risk as a result of being open about his/her feelings about his/her group's role in the implementation of Colorado Senate Bill 10-191. Again, I will strive to make sure that each participant feels safe in the group discussion.

Privacy/Confidentiality
I will protect each participant's confidentiality by transferring all real names (of participants and schools) into fictitious names in the methods chapter of my dissertation. Following this, if I were to quote any participant directly in the dissertation, I would use only the fictitious name.

I understand that there are two exceptions to the promise of confidentiality. If information is revealed concerning suicide, homicide or child abuse and neglect, it is required by law that this be reported to the proper authorities. In addition, should any information contained in this study be the subject of a court order or lawful subpoena, the University of Denver might not be able to avoid compliance with the order or subpoena.

Taking part is voluntary
As a participant, your participation is voluntary. You may refuse to participate before the study begins, discontinue at any time, or refrain from answering any questions that make you feel uncomfortable, with no penalty to you.

If you have questions
The main researcher conducting this study is Sarah Melvoin Bridich, a graduate student at the University of Denver. Please ask any questions you have now. If you have questions later, you may contact Sarah Bridich at sbridich@du.edu or 720.219.9650.

If you have any concerns or complaints about how you were treated during the interview, please contact Paul Olk, Chair, Institutional Review Board for the Protection of Human Subjects, at 303-871-4531, or du-irb@du.edu, Office of Research and Sponsored Programs at 303-871-4050 or write to either at the University of Denver, Office of Research and Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208-4820.

You will be given a copy of this form to keep for your records.

Statement of Consent
I have read and understood the foregoing descriptions of Sarah Melvoin Bridich's Dissertation research project. I have asked for and received a satisfactory explanation of any language that I did not fully understand. I agree to participate in this study, and I understand that I may withdraw my consent at any time. I have received a copy of this consent form.

Your Signature_____ Date_____

Your Name (printed)_____

Signature of person obtaining consent_____ Date_____

Printed name of person obtaining consent_____

This consent form will be kept by the researcher for at least five years beyond the end of the study.

Appendix E

Interview Questions

1. 79% of teachers and administrators (but 56% of policymakers) support the linking of student assessment data created by teachers to teacher evaluations. But 69% of teachers do NOT support the linking of statewide assessment data (such as TCAP) to teacher evaluations while 72% of administrators and 55% of policymakers are okay with it. Why do you think this is the case?

2. The majority of teachers and administrators believe both that building administrators have "primary responsibility" for the successful implementation of SB 191 and that the central focus of the SB 191 reform was teachers. How do you reconcile this?

3. Over 92% of everyone who took the survey (RSD teachers and administrators, and state policymakers) agreed that the ideal teacher evaluation system will "help teachers become better teachers." But only 59% of all teachers (compared to 76% of administrators and 78% of policymakers) agreed that an ideal teacher evaluation system will distinguish strong teachers from weak teachers. Why do you think this is?

4. The overwhelming majority of teachers, administrators, and policymakers agree that both policymakers and teachers, and policymakers and building administrators, do not have the same perceptions about what is needed for education reform. These same groups also agreed, collectively at approximately 67%, that teachers and building administrators do have similar perceptions. Yet the data from this study indicate that teachers and administrators disagree far more than teachers and policymakers or building admin and policymakers. What do you make of this finding?

5. Over 93% of everyone who took the survey agreed that ideal teacher evaluations have the potential to improve student learning. Yet only 55% of teachers, compared to 83% of administrators and 67% of policymakers, believe that SB 191 has a chance to improve student learning. What do you think may explain this gap?

6. Similarly, 55% of both policymakers and teachers believe that education reforms can improve student learning, but almost 75% of administrators believe this is possible. What do you think explains this gap in perceptions?

7. Over 60% of all administrators believe the current teacher evaluation process is both thorough and thoughtful, but only 44% of teachers think it is thorough and

55% think it is thoughtful. Why do you think this gap exists between teachers and administrators? Also, why might teachers feel that the current process is more thoughtful than thorough?

8. Almost everyone who took the survey agrees (over 93%) that there is, in fact, a right way and a wrong way to introduce new policies in a school. What do you think is the right way and/or wrong way at your school? How might this knowledge apply to the introduction of new teacher evaluations in RSD next year?

9. Similarly, over 97% of teachers "respond well to work when involved with creating the change," compared to 82% who respond well to changes at work "when informed of the new changes," which is still a large percentage. Yet only 46% of teachers believe they will play an active role in the implementation of the new teacher evaluation system in their school. What do you think of this anticipated low rate of participation for teachers in the implementation of the new teacher evaluation system? Particularly in light of the data about change preferences?

10. The following data applies to teachers:

    a. 50% stated that they did not have a strong understanding of the bill;
    b. 70% that they did not like the bill in its final version;
    c. 55% that the bill does not have a chance of improving student learning.

    What do you make of this data collection?

11. (Time permitting) Anything else you would like to add?

Appendix F

RSD & GA Survey Consent Form

What the study is about
The purpose of this research is to explore Colorado's teachers', administrators' and policymakers' perceptions of their role in the implementation of Colorado Senate Bill 10-191.

What I will ask you to do
I am asking you to participate in this study by completing a survey designed to explore individual's perspectives on many elements of teacher evaluations, education reforms, and Colorado Senate Bill 10-191. The survey should take approximately 10 minutes.

Taking part is voluntary
As a participant, your participation is voluntary. You may refuse to participate before the survey begins, discontinue at any time, or refrain from answering any questions that make you feel uncomfortable.

Potential Benefits and Risks.
This study is designed to provide the researcher with important information about teachers and administrators feel about teacher evaluations, education reforms, and Colorado Senate Bill 10-191. The researcher is focusing on teachers' and administrators' perceptions because she values their insights. Therefore, your participation in this study is greatly valued and will enable the researcher to gain insights into precisely how teachers and administrators feel in this district. Thus, you will be contributing to a study that may positively impact teachers, administrators and policymakers in the future. The risks associated with this study appear to be minimal. However, many of the questions ask you to share your personal feelings. If any question makes you uncomfortable, though, you are able to skip it or discontinue at any time.

Confidentiality
All survey information is confidential. Only the researcher will have access to the original data (i.e., the individual survey results). An identification number will be used with each survey completed, but no names will be connected to individual surveys. All the data will be kept in a secure setting. Additionally, when the researcher reports information, it will be reported for entire groups of participants and never for an individual. Every effort will be made to protect the privacy of individuals who take this survey.

I understand that there are two exceptions to the promise of confidentiality. If information is revealed concerning suicide, homicide or child abuse and neglect, it is required by law that this be reported to the proper authorities. In addition, should any information contained in this study be the subject of a court order or lawful subpoena, the University of Denver might not be able to avoid compliance with the order or subpoena.

If you have questions

The main researcher conducting this study is Sarah Melvoin Bridich, a graduate student at the University of Denver (and a former teacher and administrator). If you have questions, please email her at sbridich@du.edu. Please note that this survey was approved by the Universityof Denver's Institutional Review Board for the Protection of Human Subjects in Research on 10/04/2012.

Statement of Consent

I have read and understood the foregoing descriptions of Sarah Melvoin Bridich's Dissertation research project. I have asked for and received a satisfactory explanation of any language that I did not fully understand. By taking the survey, I agree to participate in this study, and I understand that I may withdraw my consent at any time.

Appendix G

Dissertation Research Study Consent Form (Interview)

I am asking you to participate in a research study. This form is designed to give you information about this study. I will describe this study to you and answer any of your questions.

Project Title: *Implementation Roles: Teachers', Administrators, and Education Policymakers' Perceptions of their Roles in the Implementation of Colorado Senate Bill 10-191*

Principal Investigator: Sarah M. Bridich
University of Denver, Morgridge College of Education, Education Administration and Leadership
sbridich@du.edu

Faculty Advisor: Dr. Kent Seidel
University of Denver, Morgridge College of Education, Education Administration and Leadership
kent.seidel@du.edu

What the study is about
The purpose of this research is to explore Colorado's teachers', administrators' and policymakers' perceptions of their role in the implementation of Colorado Senate Bill 10-191.

What I will ask you to do
I will ask you to participate in an interview about the topic: the implementation of Colorado Senate Bill 10-191. I will ask all participants the same questions, based on the quantitative data, and then the let the conversation go wherever is natural based on your responses to the questions. I anticipate the discussion to last approximately 20-30 minutes.

Risks and discomforts
I do not anticipate any physical, legal, or economic risks associated with this study. The conversation will be about implementation reform, particularly the new teacher evaluation system. That said, teacher evaluation can be an emotional topic and, consequently, there could be emotional risks associated with participation. Every effort will be made by the interviewer (myself) to make you feel as comfortable as possible discussing the content.

Benefits
Ideally, participants will benefit from their participation in this interview process in two ways. First, it offers individuals the chance to step back from the daily grind of actualizing reform and explore the big picture ideas behind it. Second, individuals will be contributing to a study that may benefit other teachers, administrators, and/or policymakers in the future.

Audio/Video Recording

In order to create records that best honor the words of each participant, I will make an audio recording of each interview. The recordings will then be transcribed. This means that I will have a written record of each interview that I can then read through to select ideas and quotes that best illustrate particular ideas expressed by the interviewee.

During the transcription process, all individual names and school names will be changed to protect the identities of participants.

I anticipate archiving the transcription upon completion of this project.

Please sign below if you are willing to have this interview audio recorded. Because of the research needs behind recording each interview, you may NOT participate in this study if you are not willing to be recorded.

- ☐ **I do not want to have this interview recorded; I will not participate in the study.**
- ☐ **I am willing to have this focus group recorded.**

Signed: _____

Date: _____

If the dissertation proves worthy of publications and/or presentations, I anticipate publishing and/or presenting. However, the audio-recording would not be included in any of this, only the transcript of the interview with the coded names and schools.

Privacy/Confidentiality

I will protect each participant's confidentiality by transferring all real names (of participants and schools) into fictitious names for the transcript of the interviews. Following this, if I were to quote any participant directly in the dissertation, I would use only the fictitious name.

Information about the interviews may be disseminated through emails. Please note that email communication is neither private nor secure. Though I am taking precautions to protect your privacy, you should be aware that information sent through e-mail could be read by a third party.

I understand that there are two exceptions to the promise of confidentiality. If information is revealed concerning suicide, homicide or child abuse and neglect, it is required by law that this be reported to the proper authorities. In addition, should any information contained in this study be the subject of a court order or lawful subpoena, the University of Denver might not be able to avoid compliance with the order or subpoena.

Taking part is voluntary
As a participant, your participation is voluntary. You may refuse to participate before the study begins, discontinue at any time, or refrain from answering any questions that make you feel uncomfortable, with no penalty to you.

If you have questions
The main researcher conducting this study is Sarah Melvoin Bridich, a graduate student at the University of Denver. Please ask any questions you have now. If you have questions later, you may contact Sarah Bridich at sbridich@du.edu or 720.219.9650.

If you have any concerns or complaints about how you were treated during the interview, please contact Paul Olk, Chair, Institutional Review Board for the Protection of Human Subjects, at 303-871-4531, or du-irb@du.edu, Office of Research and Sponsored Programs at 303-871-4050 or write to either at the University of Denver, Office of Research and Sponsored Programs, 2199 S. University Blvd., Denver, CO 80208-4820.

You will be given a copy of this form to keep for your records.

Statement of Consent
I have read and understood the foregoing descriptions of Sarah Melvoin Bridich's Dissertation research project. I have asked for and received a satisfactory explanation of any language that I did not fully understand. I agree to participate in this study, and I understand that I may withdraw my consent at any time. I have received a copy of this consent form.

Your Signature_____ Date_____

Your Name (printed)_____

Signature of person obtaining consent_____ Date_____

Printed name of person obtaining consent_____

This consent form will be kept by the researcher for at least five years beyond the end of the study.

Appendix H

Survey, by Analysis Label of 40 Likert Questions

	Label	Question from Survey
1	C-TE_1	Current TE improved your teaching.
2	C-TE_2	Current TE has provided me with meaningful feedback.
3	C-TE_3	Current TE has made me a better teacher overall.
4	C-TE_1a	Current TE encourages professional growth for teachers.
5	C-TE_2a	Current TE is linked to student learning.
6	C-TE1_3a	Current TE is able to assess teachers' overall teaching abilities.
7	C-TE_4a	Current TE is a high stress process for teachers.
8	C-TE_5a	Current TE is a high stress process for administrators.
9	C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.
10	C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.
11	IdealTE_1	Ideal TE distinguish strong teachers from weak teachers.
12	IdealTE_2	Ideal TE help teachers become better teachers.
13	IdealTE_3	Ideal TE have the ability to improve student learning.
14	IdealTE_4	Ideal TE are linked to student assessment data of any form.
15	IdealTE_5	Ideal TE are linked to student assessment data by the state department of education (i.e., CSAPs).
16	IdealTE_6	Ideal TE are linked to student assessment data created by the teacher.
17	Change_1	I view change as an opportunity for growth.
18	Change_2	I respond well when policies change at work.
19	Change_3	I respond well to change at work when I am involved with creating the change.
20	Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.
21	Change_5	There is a right way and a wrong way to introduce new policies in a school.

22	EdRef_1	EdRef: Education reforms improve student learning.
23	EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.
24	EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.
25	EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.
26	EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public ed.
27	EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.
28	SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.
29	SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.
30	SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.
31	PrimaryResp	Which group has primary responsibility for successful implementation of new teacher evaluations?
32	SB191_1	SB 191: I have a strong understanding of the legislation SB 191.
33	SB191_2	SB 191: I was highly involved in the development of SB 191 during its implementation phase.
34	SB191_3	SB 191: I supported SB 191 during its development.
35	SB191_4	SB 191: I was pleased with the final version of SB 191 as it was passed.
36	SB191_5	SB 191: The media portrayed teachers fairly during the bill's passage.
37	SB191_6	SB 191: The media portrayed administrators fairly during the bill's passage.
38	SB191_7	SB 191: The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.
39	SB191_8	SB 191: The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.
40	CntrlFocus	Central focus of SB 191

Appendix I

Ethnic Breakdown of Survey Participants

Table

Ethnicity of Participants

	Asian	African American	Hispanic/ Latino	Native American or Alaskan Native	Native Hawaiian or Pacific Islander	White	Sub-totals	missing
Teachers	7	13	21	8	2	476	527	62
Administrators	0	7	8	0	0	39	54	10
Policymakers	0	1	0	0	0	8	9	3
Totals	7	21	29	8	2	523	590	75

Note: The totals listed in the bottom row include all who checked that ethnic box; however, many of those who checked an ethnicity other than "white" checked multiple boxes, resulting in numbers that are higher than those in the "minority" and "white" categories in Table 3.

Appendix J

Pearson Chi-Square Test Results for Teachers by School Level Subgroup

Table J1

Current Teacher Evaluations, Teachers by School Level

Label	Question:	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	71%	66%	65%	65%	9.367	9	.404
C-TE_2	Current TE has provided me with meaningful feedback.	73%	68%	68%	65%	11.472	9	.245
C-TE_3	Current TE has made me a better teacher overall.	68%	61%	63%	65%	9.478	9	.394
C-TE_1a	Current TE encourages professional growth for teachers.	74%	60%	68%	70%	10.589	9	.305
C-TE_2a	Current TE is linked to student learning.	76%	64%	65%	65%	21.668	9	.010
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	66%	57%	57%	65%	15.120	9	.088
C-TE_4a	Current TE is a high stress process for teachers.	69%	61%	58%	75%	26.127	9	.002
C-TE_5a	Current TE is a high stress process for administrators.	62%	49%	39%	55%	24.949	9	.003
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	50%	40%	35%	45%	16.522	9	.057
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	59%	53%	50%	53%	18.237	9	.033
	Counts	235	114	144	20	–	–	–

Note: Counts are an average. For all questions, at least two cells had a count less than five, which may have impacted significance.

Table J2

Ideal Teacher Evaluations, Teachers by School Level

Label	Question: Ideal teacher evaluations:	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	57%	63%	56%	79%	9.532	9	.390
IdealTE_2	help teachers become better teachers.	93%	95%	95%	90%	14.366	9	.111
IdealTE_3	have the ability to improve student learning.	92%	95%	92%	95%	8.841	9	.452
IdealTE_4	are linked to student assessment data of any form.	67%	68%	62%	60%	8.087	9	.525
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	30%	39%	24%	30%	20.078	9	.017
IdealTE_6	are linked to student assessment data created by the teacher.	78%	82%	76%	85%	6.949	9	.642
	Counts	237	114	144	20	–	–	–

Note: Counts are an average. For all questions, at least one cell had a count less than five, which may have impacted significance.

Table J3

Education Reform, Teachers by School Level

Label	Question	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	62%	57%	38%	70%	32.965	9	.000
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	41%	32%	23%	45%	22.506	9	.007
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	54%	59%	52%	65%	9.515	9	.391
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	6%	1%	3%	0%	13.254	6	.039
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	7%	5%	6%	5%	10.880	9	.284
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	76%	62%	60%	65%	21.463	9	.011
	Counts	236	114	144	20	–	–	–

Note: Counts are an average. For all questions, at least two cells had a count less than five, which may have impacted significance.

Table J4

SB 191 Teacher Evaluations, Teachers by School Level

Label	Question	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	50%	39%	45%	45%	26.601	12	.009
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	14%	16%	12%	5%	38.904	12	.000
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	13%	18%	17%	10%	38.174	12	.000
	Counts	236	114	145	20	–	–	–

Note: Counts are an average for elementary school teachers. For all questions, at least three cells had a count less than five, which may have impacted significance.

Table J5

SB191 General, Teachers by School Level

Label	Question	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	43%	44%	55%	45%	14.956	12	.244
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	3%	4%	2%	5%	5.437	12	.942
SB191_3	I supported SB 191 during its development.	15%	18%	10%	5%	31.956	12	.001
SB191_4	I was pleased with the final version of SB 191 as it was passed.	15%	16%	11%	10%	38.714	12	.000
SB191_5	The media portrayed teachers fairly during the bill's passage.	10%	11%	6%	5%	20.138	12	.065
SB191_6	The media portrayed administrators fairly during the bill's passage.	14%	19%	15%	20%	12.437	12	.411
SB191_7	The media portrayed policymakers (i.e., legislators, lobbyists) fairly during the bill's passage.	35%	34%	28%	25%	13.552	12	.330
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	38%	41%	30%	35%	27.487	12	.007
	Counts	235	113	145	20	–	–	–

Note: Counts are an average. For all questions, at least three cells had a count less than five, which may have impacted significance.

Table J6

Change Questions, Teachers by School Level

Label	Question	Elem.	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
Chg_1	I view change as an opportunity for growth.	97%	97%	92%	95%	14.845	9	.095
Chg_2	I respond well when policies change at work.	78%	80%	72%	85%	5.775	9	.762
Chg_3	I respond well to change at work when I am involved with creating the change.	98%	98%	97%	100%	8.053	6	.234
Chg_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	83%	86%	77%	75%	9.904	9	.358
Chg_5	There is a right way and a wrong way to introduce new policies in a school.	92%	97%	97%	90%	10.605	9	.304
	Counts	237	113	145	20	–	–	–

Note: Counts are an average for middle school teachers. For all questions, at least three cells had an expected count less than five, which may have impacted significance.

Appendix K

Pearson Chi-Square Test Results for Teachers by Gender Subgroup

Table K1

Current Teacher Evaluations, Teachers by Gender

Label	Question:	Male	Female	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	64%	69%	3.397	3	.334
C-TE_2	Current TE has provided me with meaningful feedback.	72%	70%	4.360	3	.225
C-TE_3	Current TE has made me a better teacher overall.	63%	65%	6.908	3	.075
C-TE_1a	Current TE encourages professional growth for teachers.	63%	70%	3.412	3	.332
C-TE_2a	Current TE is linked to student learning.	65%	70%	4.482	3	.214
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	57%	63%	3.653	3	.301
C-TE_4a	Current TE is a high stress process for teachers.	60%	65%	2.525	3	.471
C-TE_5a	Current TE is a high stress process for administrators.	50%	54%	1.446	3	.695
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	35%	46%	6.220	3	.101
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	48%	56%	4.783	3	.188
	Counts	107	403	–	–	–

Note. Counts are an average. C-TE_4 and C-TE_6 had at least one cell with a count less than 5, which may have impacted significance.

Table K2

Ideal Teacher Evaluations, Teachers by Gender

Label	Question: Ideal teacher evaluations:	Male	Female	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	56%	60%	4.119	3	.249
IdealTE_2	help teachers become better teachers.	93%	94%	5.053	3	.168
IdealTE_3	have the ability to improve student learning.	89%	93%	2.712	3	.438
IdealTE_4	are linked to student assessment data of any form.	55%	69%	9.676	3	.022
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	27%	31%	5.895	3	.117
IdealTE_6	are linked to student assessment data created by the teacher.	74%	80%	1.643	3	.650
	Counts	108	406	–	–	–

Note. Counts are an avearge. IdealTE_2, IdealTE_3, IdealTE_5, and IdealTE_6 had at least one cell with an expected count less than 5, which may have impacted significance.

Table K3

Education Reform, Teachers by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	47%	57%	8.421	3	.038
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	34%	34%	5.475	3	.140
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	56%	55%	0.281	3	.964
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	3%	4%	0.767	2	.681
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	5%	7%	2.442	3	.486
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	63%	70%	2.474	3	.480
	Counts	108	405	–	–	–

Note. Counts are an avearge. All questions except EdRef_6 had at least one cell with an expected count less than 5, which may have impacted significance.

Table K4

SB191 Teacher Evaluations, Teachers by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	47%	46%	1.272	4	.866
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	14%	13%	1.082	4	.897
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	19%	14%	4.694	4	.320
	Counts	108	405	–	–	–

Note. Counts an average for female teachers. SB191TE_2 and SB191TE_3 had at least one cell with an expected count less than 5, which may have impacted significance.

Table K5

SB191 General, Teachers by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	59%	44%	22.047	4	.000
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	4%	3%	6.157	4	.188
SB191_3	I supported SB 191 during its development.	16%	14%	4.739	4	.315
SB191_4	I was pleased with the final version of SB 191 as it was passed.	19%	13%	8.318	4	.081
SB191_5	The media portrayed teachers fairly during the bill's passage.	9%	9%	7.060	4	.133
SB191_6	The media portrayed administrators fairly during the bill's passage.	18%	15%	2.066	4	.724
SB191_7	The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.	36%	31%	4.183	4	.382
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	34%	37%	6.277	4	.179
	Counts	108	402	–	–	–

Note. Counts are an average. All questions except SB191_7 and SB191_8 had at least one cell with an expected count less than 5, which may have impacted significance.

Table K6

Change, Teachers by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	94%	96%	8.484	3	.037
Change_2	I respond well when policies change at work.	69%	80%	5.519	3	.137
Change_3	I respond well to change at work when I am involved with creating the change.	97%	98%	0.628	2	.731
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	75%	84%	8.580	3	.035
Change_5	There is a right way and a wrong way to introduce new policies in a school.	98%	93%	3.771	3	.287
	Counts	108	405	–	–	–

Note. Counts are an average for female teachers. All questions had at least one cell with an expected count less than 5, which may have impacted significance.

Appendix L

Pearson Chi-Square Test Results for Teachers by Educator Type Subgroup

Table L1

Current Teacher Evaluations, Teachers by Educator Type

Label	Question:	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	72%	68%	63%	8.345	6	.214
C-TE_2	Current TE has provided me with meaningful feedback.	73%	72%	64%	14.114	6	.028
C-TE_3	Current TE has made me a better teacher overall.	68%	65%	59%	10.479	6	.106
C-TE_1a	Current TE encourages professional growth for teachers.	70%	69%	70%	3.590	6	.732
C-TE_2a	Current TE is linked to student learning.	70%	70%	72%	0.665	6	.995
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	64%	60%	62%	4.344	6	.630
C-TE_4a	Current TE is a high stress process for teachers.	63%	64%	65%	6.233	6	.398
C-TE_5a	Current TE is a high stress process for administrators.	52%	49%	57%	10.218	6	.116
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	42%	44%	49%	4.462	6	.614
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	54%	54%	59%	4.021	6	.674
	Counts	265	153	132	—	—	—

Note. Counts are an average. C-TE_4a and C-TE_6a had at least one cell with an expected count less than five, which may have impacted significance.

Table L2

Ideal Teacher Evaluations, Teachers by Educator Type

Label	Question: Ideal teacher evaluations:	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	62%	48%	66%	18.495	6	.005
IdealTE_2	help teachers become better teachers.	95%	92%	94%	5.729	6	.454
IdealTE_3	have the ability to improve student learning.	94%	90%	91%	4.274	6	.640
IdealTE_4	are linked to student assessment data of any form.	70%	61%	65%	6.076	6	.415
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	36%	23%	32%	12.242	6	.057
IdealTE_6	are linked to student assessment data created by the teacher.	80%	76%	82%	11.414	6	.076
	Counts	263	151	127	—	—	—

Note. Counts are an average. IdealTE_2 and IdealTE_3 had at least three cells with an expected count less than five, which may have impacted significance.

Table L3

Education Reform, Teachers by Educator Type

Label	Question	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	54%	51%	62%	9.412	6	.152
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	34%	36%	35%	5.777	6	.449
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	58%	53%	52%	9.690	6	.138
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	4%	3%	7%	5.780	4	.216
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	5%	9%	8%	11.244	6	.081
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	72%	63%	68%	9.112	6	.167
	Counts	256	152	122	–	–	–

Note. Counts are an average. All questions, excpet EdRef_4 and EdRef_6, had at least two cells with an expected count less than five, which may have impacted significance.

Table L4

SB191 Teacher Evaluations, Teachers by Educator Type

Label	Question	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	46%	51%	41%	6.856	8	.552
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	13%	11%	16%	9.173	8	.328
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	15%	14%	18%	7.593	8	.474
	Counts	255	150	122	–	–	–

Note. Counts are an average. SB191TE_2 and SB191TE_3 had at least three cells with an expected count less than five, which may have impacted significance.

Table L5

SB191, Teachers by Educator Type

Label	Question	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	46%	54%	40%	13.559	8	.094
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	3%	4%	3%	4.427	8	.817
SB191_3	I supported SB 191 during its development.	15%	15%	11%	7.471	8	.487
SB191_4	I was pleased with the final version of SB 191 as it was passed.	13%	16%	14%	2.733	8	.950
SB191_5	The media portrayed teachers fairly during the bill's passage.	9%	10%	8%	5.954	8	.652
SB191_6	The media portrayed administrators fairly during the bill's passage.	16%	18%	13%	5.197	8	.736
SB191_7	The media portrayed policymakers (i.e., legislators, lobbyists) fairly during the bill's passage.	36%	27%	30%	7.738	8	.459
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	38%	34%	36%	12.038	8	.149
	Counts	252	145	120	–	–	–

Note. Counts are an average. All questions, except SB101_7 and SB191_8, had at least three cells with an expected+A1 count less than five, which may have impacted significance.

Table L6

Change, Teachers by Educator Type

Label	Question	TCAP-subject Teacher	non-TCAP subject Teacher	Other Licensed Professional	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	95%	95%	98%	3.169	4	.530
Change_2	I respond well when policies change at work.	78%	72%	81%	9.584	6	.143
Change_3	I respond well to change at work when I am involved with creating the change.	98%	97%	98%	12.872	4	.012
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	83%	82%	79%	5.272	6	.509
Change_5	There is a right way and a wrong way to introduce new policies in a school.	95%	92%	93%	6.884	6	.332
	Counts	260	152	126	–	–	–

Note. Counts are an average. All questions, except Change_1, had at least one cell with an expected count less than five, which may have impacted significance.

Appendix M

Pearson Chi-Square Test Results for Teachers by License Type Subgroup

Table M1

Current Teacher Evaluations, Teachers by License Type

Label	Question:	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	85%	68%	6.351	3	.096
C-TE_2	Current TE has provided me with meaningful feedback.	78%	70%	11.130	3	.011
C-TE_3	Current TE has made me a better teacher overall.	78%	64%	5.188	3	.159
C-TE_1a	Current TE encourages professional growth for teachers.	73%	69%	1.865	3	.601
C-TE_2a	Current TE is linked to student learning.	75%	69%	0.872	3	.832
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	68%	61%	1.611	3	.657
C-TE_4a	Current TE is a high stress process for teachers.	58%	65%	2.488	3	.477
C-TE_5a	Current TE is a high stress process for administrators.	33%	54%	11.256	3	.010
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	58%	42%	4.844	3	.184
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	65%	54%	2.050	3	.562
	Counts	40	471	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table M2

Ideal Teacher Evaluations, Teachers by License Type

Label	Question: Ideal teacher evaluations:	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	53%	59%	2.982	3	.394
IdealTE_2	help teachers become better teachers.	95%	94%	0.887	3	.892
IdealTE_3	have the ability to improve student learning.	90%	93%	0.914	3	.822
IdealTE_4	are linked to student assessment data of any form.	77%	65%	2.897	3	.408
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	35%	30%	1.497	3	.683
IdealTE_6	are linked to student assessment data created by the teacher.	80%	78%	0.402	3	.940
	Counts	40	472	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table M3

Education Reform, Teachers by License Type

Label	Question	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	73%	53%	7.281	3	.063
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	55%	32%	8.602	3	.035
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	70%	54%	5.557	3	.135
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	8%	3%	3.030	2	.220
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	15%	6%	6.669	3	.083
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	70%	68%	0.936	3	.817
	Counts	40	471	–	–	–

Note. Counts are an average for teachers with Professional Licenses. All questions had at least one cell with an expected less than five, which may have impacted significance.

Table M4

SB 191 Teacher Evaluations, Teachers by License Type

Label	Question	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	43%	46%	1.453	4	.835
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	38%	11%	25.531	4	.000
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	35%	13%	13.927	4	.008
	Counts	40	472	–	–	–

Note. Counts are an average for teachers with Professional Licenses. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table M5

SB191, Teachers by License Type

Label	Question	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	38%	47%	3.870	4	.424
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	3%	3%	2.961	4	.564
SB191_3	I supported SB 191 during its development.	18%	14%	5.671	4	.225
SB191_4	I was pleased with the final version of SB 191 as it was passed.	28%	13%	10.524	4	.032
SB191_5	The media portrayed teachers fairly during the bill's passage.	33%	7%	49.321	4	.000
SB191_6	The media portrayed administrators fairly during the bill's passage.	33%	14%	25.678	4	.000
SB191_7	The media portrayed policymakers (i.e., legislators, lobbyists) fairly during the bill's passage.	43%	32%	5.937	4	.204
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	45%	36%	17.164	4	.002
	Counts	40	469	–	–	–

Note. Counts are an average for teachers with Professional Licenses. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table M6

Change, Teachers by License Type

Label	Question	Non-Professional	Professional	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	98%	95%	0.744	3	.863
Change_2	I respond well when policies change at work.	70%	78%	18.228	3	.000
Change_3	I respond well to change at work when I am involved with creating the change.	100%	98%	2.083	2	.353
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	80%	81%	7.468	3	.058
Change_5	There is a right way and a wrong way to introduce new policies in a school.	93%	95%	0.702	3	.873
	Counts	40	472	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Appendix N

Pearson Chi-Square Test Results for Teachers by Ethnicity Subgroup

Table N1

Current Teacher Evaluations, Teachers by Ethnicity

Label	Question:	Minority	White	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	67%	68%	2.690	3	.442
C-TE_2	Current TE has provided me with meaningful feedback.	66%	70%	1.938	3	.585
C-TE_3	Current TE has made me a better teacher overall.	70%	64%	1.259	3	.739
C-TE_1a	Current TE encourages professional growth for teachers.	71%	69%	0.518	3	.915
C-TE_2a	Current TE is linked to student learning.	76%	69%	2.920	3	.404
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	64%	61%	1.092	3	.779
C-TE_4a	Current TE is a high stress process for teachers.	69%	64%	1.313	3	.726
C-TE_5a	Current TE is a high stress process for administrators.	65%	52%	4.114	3	.249
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	49%	43%	1.476	3	.688
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	62%	54%	3.237	3	.357
	Counts	45	465	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table N2

Ideal Teacher Evaluations, Teachers by Ethnicity

Label	Question: Ideal teacher evaluations:	Minority	White	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	37%	61%	13.733	3	.003
IdealTE_2	help teachers become better teachers.	93%	94%	10.005	3	.019
IdealTE_3	have the ability to improve student learning.	91%	92%	6.005	3	.111
IdealTE_4	are linked to student assessment data of any form.	67%	66%	0.935	3	.817
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	22%	31%	1.948	3	.583
IdealTE_6	are linked to student assessment data created by the teacher.	82%	79%	1.320	3	.724
	Counts	45	467	–	–	–

Note. Counts are an average. All questions, except IdealTE_1, had at least one cell with an expected count less than five, which may have impacted significance.

Table N3

Education Reform, Teachers by Ethnicity

Label	Question	Minority	White	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	62%	54%	5.661	3	.129
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	32%	34%	0.821	3	.844
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	64%	54%	3.809	3	.283
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	11%	3%	8.697	2	.013
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	9%	6%	1.237	3	.744
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	69%	68%	5.103	3	.164
	Counts	45	466	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table N4

SB191 Teacher Evaluations, Teachers by Ethnicity

Label	Question	Minority	White	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	49%	45%	2.101	4	.717
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	24%	12%	5.825	4	.213
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	22%	14%	4.493	4	.343
	Counts	45	466	–	–	–

Note. Counts are an average. All questions had one cell with an expected count less than five, which may have impacted significance.

Table N5

SB191, Teachers by Ethnicity

Label	Question	Minority	White	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	49%	46%	2.381	4	.666
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	4%	3%	5.785	4	.216
SB191_3	I supported SB 191 during its development.	18%	14%	3.604	4	.462
SB191_4	I was pleased with the final version of SB 191 as it was passed.	16%	14%	0.546	4	.969
SB191_5	The media portrayed teachers fairly during the bill's passage.	13%	8%	10.867	4	.028
SB191_6	The media portrayed administrators fairly during the bill's passage.	31%	14%	15.160	4	.004
SB191_7	The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.	40%	31%	3.636	4	.458
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	31%	37%	2.583	4	.630
	Counts	45	464	–	–	–

Note. Counts are an average. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table N6

Change, Teachers by Ethnicity

Label	Question	Minority	White	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	100%	95%	2.309	3	.511
Change_2	I respond well when policies change at work.	78%	78%	0.936	3	.817
Change_3	I respond well to change at work when I am involved with creating the change.	98%	98%	0.052	2	.974
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	89%	81%	3.107	3	.375
Change_5	There is a right way and a wrong way to introduce new policies in a school.	93%	94%	0.382	3	.944
	Counts	45	467	–	–	–

Note. Counts are an average for white teachers. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Appendix O

One-Way ANOVA with post-hoc Scheffé Test Results for Teachers by School Level Subgroup

Table O1

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 5.537, p = .001

Current TE is linked to student learning. (CTE_2a)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.264*	.085	.023	.03	.50
	High School	.282*	.079	.006	.06	.50
	K-8 + Other	.157	.174	.845	-.33	.65
Middle School	Elementary School	-.264*	.085	.023	-.50	-.03
	High School	.017	.094	.998	-.25	.28
	K-8 + Other	-.107	.181	.951	-.62	.40
High School	Elementary School	-.282*	.079	.006	-.50	-.06
	Middle School	-.017	.094	.998	-.28	.25
	K-8 + Other	-.124	.178	.922	-.62	.38
K-8 + Other	Elementary School	-.157	.174	.845	-.65	.33
	Middle School	.107	.181	.951	-.40	.62
	High School	.124	.178	.922	-.38	.62

*. The mean difference is significant at the 0.05 level.

Table O2

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 2.505 p = .058

Current TE is a high stress process for teachers. (CTE_4a)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.185	.098	.312	-.09	.46
	High School	.219	.091	.120	-.03	.47
	K-8 + Other	.233	.199	.714	-.33	.79
Middle School	Elementary School	-.185	.098	.312	-.46	.09
	High School	.034	.107	.992	-.27	.34
	K-8 + Other	.048	.208	.997	-.53	.63
High School	Elementary School	-.219	.091	.120	-.47	.03
	Middle School	-.034	.107	.992	-.34	.27
	K-8 + Other	.014	.204	1.000	-.56	.59
K-8 + Other	Elementary School	-.233	.199	.714	-.79	.33
	Middle School	-.048	.208	.997	-.63	.53
	High School	-.014	.204	1.000	-.59	.56

Table O3

One-Way ANOVA with Scheffé Post-Hoc Test

F (5, 498) = 6.123, p = .000

Current TE is a high stress process for administrators. (CTE_5a)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.210	.084	.102	-.03	.45
	High School	.320*	.078	.001	.10	.54
	K-8 + Other	.251	.170	.537	-.23	.73
Middle School	Elementary School	-.210	.084	.102	-.45	.03
	High School	.110	.093	.706	-.15	.37
	K-8 + Other	.041	.177	.997	-.46	.54
High School	Elementary School	-.320*	.078	.001	-.54	-.10
	Middle School	-.110	.093	.706	-.37	.15
	K-8 + Other	-.069	.175	.985	-.56	.42
K-8 + Other	Elementary School	-.251	.170	.537	-.73	.23
	Middle School	-.041	.177	.997	-.54	.46
	High School	.069	.175	.985	-.42	.56

*. The mean difference is significant at the 0.05 level.

Table O4

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 3.011 p = .030

Current TE allows for thoughtful reviews of teachers' overall teaching abilities. (CTE_7a)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.180	.090	.267	-.07	.43
	High School	.198	.084	.134	-.04	.43
	K-8 + Other	.355	.189	.318	-.17	.89
Middle School	Elementary School	-.180	.090	.267	-.43	.07
	High School	.018	.099	.998	-.26	.30
	K-8 + Other	.175	.197	.850	-.38	.73
High School	Elementary School	-.198	.084	.134	-.43	.04
	Middle School	-.018	.099	.998	-.30	.26
	K-8 + Other	.158	.193	.882	-.39	.70
K-8 + Other	Elementary School	-.355	.189	.318	-.89	.17
	Middle School	-.175	.197	.850	-.73	.38
	High School	-.158	.193	.882	-.70	.39

Table O5

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 4.249 p = .006

Ideal TE are linked to student assessment data by the state department of education (i.e., CSAPs). (IdealTE_5)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.110	.093	.707	-.37	.15
	High School	.233	.086	.065	-.01	.48
	K-8 + Other	-.064	.190	.990	-.60	.47
Middle School	Elementary School	.110	.093	.707	-.15	.37
	High School	.343*	.102	.011	.06	.63
	K-8 + Other	.046	.198	.997	-.51	.60
High School	Elementary School	-.233	.086	.065	-.48	.01
	Middle School	-.343*	.102	.011	-.63	-.06
	K-8 + Other	-.297	.195	.509	-.84	.25
K-8 + Other	Elementary School	.064	.190	.990	-.47	.60
	Middle School	-.046	.198	.997	-.60	.51
	High School	.297	.195	.509	-.25	.84

*. The mean difference is significant at the 0.05 level.

Table O6

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 509) = 10.613, p = .000

Education reforms improve student learning. (EdRef_1)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.105	.070	.521	-.09	.30
	High School	.351*	.064	.000	.17	.53
	K-8 + Other	-.064	.142	.976	-.46	.33
Middle School	Elementary School	-.105	.070	.521	-.30	.09
	High School	.246*	.076	.016	.03	.46
	K-8 + Other	-.169	.148	.726	-.58	.24
High School	Elementary School	-.351*	.064	.000	-.53	-.17
	Middle School	-.246*	.076	.016	-.46	-.03
	K-8 + Other	-.415*	.145	.043	-.82	-.01
K-8 + Other	Elementary School	.064	.142	.976	-.33	.46
	Middle School	.169	.148	.726	-.24	.58
	High School	.415*	.145	.043	.01	.82

*. The mean difference is significant at the 0.05 level.

Table O7

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 509) = 6.010, p = .000

Education reforms view teachers as knowledgeable professionals. (EdRef_2)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.086	.084	.785	-.15	.32
	High School	.294*	.077	.003	.08	.51
	K-8 + Other	-.204	.170	.697	-.68	.27
Middle School	Elementary School	-.086	.084	.785	-.32	.15
	High School	.208	.092	.165	-.05	.47
	K-8 + Other	-.291	.177	.444	-.79	.21
High School	Elementary School	-.294*	.077	.003	-.51	-.08
	Middle School	-.208	.092	.165	-.47	.05
	K-8 + Other	-.499*	.175	.044	-.99	-.01
K-8 + Other	Elementary School	.204	.170	.697	-.27	.68
	Middle School	.291	.177	.444	-.21	.79
	High School	.499*	.175	.044	.01	.99

*. The mean difference is significant at the 0.05 level.

Table O8

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 511) = .748, p = .524

Policymakers, teachers have the same perceptions about what is needed to improve public education. (EdRef_4)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.011	.065	.999	-.19	.17
	High School	.028	.060	.974	-.14	.20
	K-8 + Other	-.172	.132	.636	-.54	.20
Middle School	Elementary School	.011	.065	.999	-.17	.19
	High School	.039	.071	.961	-.16	.24
	K-8 + Other	-.161	.137	.709	-.55	.22
High School	Elementary School	-.028	.060	.974	-.20	.14
	Middle School	-.039	.071	.961	-.24	.16
	K-8 + Other	-.200	.135	.533	-.58	.18
K-8 + Other	Elementary School	.172	.132	.636	-.20	.54
	Middle School	.161	.137	.709	-.22	.55
	High School	.200	.135	.533	-.18	.58

Table O9

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 511) = 5.898, p = .001

Teachers, building admins have the same perceptions about what is needed to improve public education. (EdRef_6)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.205	.074	.053	.00	.41
	High School	.268*	.068	.002	.08	.46
	K-8 + Other	.110	.151	.911	-.31	.53
Middle School	Elementary School	-.205	.074	.053	-.41	.00
	High School	.064	.081	.893	-.16	.29
	K-8 + Other	-.095	.157	.947	-.53	.34
High School	Elementary School	-.268*	.068	.002	-.46	-.08
	Middle School	-.064	.081	.893	-.29	.16
	K-8 + Other	-.158	.154	.788	-.59	.27
K-8 + Other	Elementary School	-.110	.151	.911	-.53	.31
	Middle School	.095	.157	.947	-.34	.53
	High School	.158	.154	.788	-.27	.59

*. The mean difference is significant at the 0.05 level.

Table O10

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 511) = 4.005, p = .008

I will play an active role in the implementation of my school's new teacher evaluation system. (SB191TE_1)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.386	.146	.074	-.80	.02
	High School	.151	.135	.743	-.23	.53
	K-8 + Other	.104	.298	.989	-.73	.94
Middle School	Elementary School	.386	.146	.074	-.02	.80
	High School	.537*	.160	.011	.09	.99
	K-8 + Other	.490	.311	.477	-.38	1.36
High School	Elementary School	-.151	.135	.743	-.53	.23
	Middle School	-.537*	.160	.011	-.99	-.09
	K-8 + Other	-.047	.306	.999	-.90	.81
K-8 + Other	Elementary School	-.104	.298	.989	-.94	.73
	Middle School	-.490	.311	.477	-1.36	.38
	High School	.047	.306	.999	-.81	.90

*. The mean difference is significant at the 0.05 level.

Table O11

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 512) = 10.324, p = .000

Teachers at my school believe it will be beneficial to the teachers' professional practice at my school. (SB191TE_2)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.492*	.162	.028	-.95	-.04
	High School	.377	.150	.100	-.04	.80
	K-8 + Other	-.880	.332	.072	-1.81	.05
Middle School	Elementary School	.492*	.162	.028	.04	.95
	High School	.868*	.178	.000	.37	1.37
	K-8 + Other	-.389	.346	.738	-1.36	.58
High School	Elementary School	-.377	.150	.100	-.80	.04
	Middle School	-.868*	.178	.000	-1.37	-.37
	K-8 + Other	-1.257*	.340	.004	-2.21	-.30
K-8 + Other	Elementary School	.880	.332	.072	-.05	1.81
	Middle School	.389	.346	.738	-.58	1.36
	High School	1.257*	.340	.004	.30	2.21

*. The mean difference is significant at the 0.05 level.

Table O12

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 10.227, p = .000

Teachers at my school believe it will be beneficial to students at my school. (SB191TE_3)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.465*	.165	.047	-.93	.00
	High School	.434*	.152	.045	.01	.86
	K-8 + Other	-.771	.336	.154	-1.71	.17
Middle School	Elementary School	.465*	.165	.047	.00	.93
	High School	.899*	.181	.000	.39	1.41
	K-8 + Other	-.306	.350	.857	-1.29	.67
High School	Elementary School	-.434*	.152	.045	-.86	-.01
	Middle School	-.899*	.181	.000	-1.41	-.39
	K-8 + Other	-1.205*	.344	.007	-2.17	-.24
K-8 + Other	Elementary School	.771	.336	.154	-.17	1.71
	Middle School	.306	.350	.857	-.67	1.29
	High School	1.205*	.344	.007	.24	2.17

*. The mean difference is significant at the 0.05 level.

Table O13

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 510) = 7.613, p = .000

I supported SB 191 during its development. (SB191_3)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.041	.160	.996	-.49	.41
	High School	.555*	.148	.003	.14	.97
	K-8 + Other	-.622	.326	.304	-1.54	.29
Middle School	Elementary School	.041	.160	.996	-.41	.49
	High School	.596*	.175	.010	.10	1.09
	K-8 + Other	-.582	.340	.403	-1.53	.37
High School	Elementary School	-.555*	.148	.003	-.97	-.14
	Middle School	-.596*	.175	.010	-1.09	-.10
	K-8 + Other	-1.178*	.334	.006	-2.11	-.24
K-8 + Other	Elementary School	.622	.326	.304	-.29	1.54
	Middle School	.582	.340	.403	-.37	1.53
	High School	1.178*	.334	.006	.24	2.11

*. The mean difference is significant at the 0.05 level.

Table O14

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 509) = 11.060, p = .000

I was pleased with the final version of SB 191 as it was passed. (SB191_4)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.092	.155	.950	-.53	.34
	High School	.624*	.143	.000	.22	1.02
	K-8 + Other	-.760	.315	.122	-1.64	.12
Middle School	Elementary School	.092	.155	.950	-.34	.53
	High School	.717*	.170	.001	.24	1.19
	K-8 + Other	-.668	.329	.249	-1.59	.25
High School	Elementary School	-.624*	.143	.000	-1.02	-.22
	Middle School	-.717*	.170	.001	-1.19	-.24
	K-8 + Other	-1.384*	.323	.000	-2.29	-.48
K-8 + Other	Elementary School	.760	.315	.122	-.12	1.64
	Middle School	.668	.329	.249	-.25	1.59
	High School	1.384*	.323	.000	.48	2.29

*. The mean difference is significant at the 0.05 level.

Table O15

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 507) = 5.918, p = .001

The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way. (SB191_8)

(I) Teachers by School Level	(J) Teachers by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.161	.132	.684	-.53	.21
	High School	.294	.121	.116	-.04	.63
	K-8 + Other	-.635	.266	.129	-1.38	.11
Middle School	Elementary School	.161	.132	.684	-.21	.53
	High School	.455*	.144	.020	.05	.86
	K-8 + Other	-.474	.278	.405	-1.25	.30
High School	Elementary School	-.294	.121	.116	-.63	.04
	Middle School	-.455*	.144	.020	-.86	-.05
	K-8 + Other	-.929*	.273	.009	-1.69	-.16
K-8 + Other	Elementary School	.635	.266	.129	-.11	1.38
	Middle School	.474	.278	.405	-.30	1.25
	High School	.929*	.273	.009	.16	1.69

*. The mean difference is significant at the 0.05 level.

Appendix P

One-Way ANOVA with post-hoc Scheffé Test Results for Teachers by Educator

Table P1

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 558) = 4.881, p = .008

Current TE has provided me with meaningful feedback. (C-TE_2)

(I) Type of Teacher/ Other Licensed Prof	(J) Type of Teacher/ Other Licensed Prof Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher of TCAP-tested subject	Teacher of non TCAP-tested subject	.011	.077	.990	-.18	.20
	Other Licensed Professional	.238*	.080	.013	.04	.43
Teacher of non TCAP-tested subject	Teacher of TCAP-tested subject	-.011	.077	.990	-.20	.18
	Other Licensed Professional	.227*	.089	.040	.01	.45
Other Licensed Professional	Teacher of TCAP-tested subject	-.238*	.080	.013	-.43	-.04
	Teacher of non TCAP-tested subject	-.227*	.089	.040	-.45	-.01

* The mean difference is significant at the 0.05 level

Table P2

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 534) = 7.157, p = .001

Ideal TE distinguish strong teachers from weak teachers. (IdealTE_1)

(I) Type of Teacher/ Other Licensed Prof	(J) Type of Teacher/ Other Licensed Prof Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher of TCAP-tested subject	Teacher of non TCAP-tested subject	.290*	.091	.007	.07	.51
	Other Licensed Professional	-.082	.097	.699	-.32	.16
Teacher of non TCAP-tested subject	Teacher of TCAP-tested subject	-.290*	.091	.007	-.51	-.07
	Other Licensed Professional	-.372*	.108	.003	-.64	-.11
Other Licensed Professional	Teacher of TCAP-tested subject	.082	.097	.699	-.16	.32
	Teacher of non TCAP-tested subject	.372*	.108	.003	.11	.64

* The mean difference is significant at the 0.05 level

Table P3

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 535) = .445, p = .641

I respond well to change at work when informed of what new policies will be and I must learn the new policies. (Change_3)

(I) Type of Teacher/ Other Licensed Prof	(J) Type of Teacher/ Other Licensed Prof Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher of TCAP-tested subject	Teacher of non TCAP-tested subject	.041	.070	.840	-.13	.21
	Other Licensed Professional	.066	.074	.670	-.12	.25
Teacher of non TCAP-tested subject	Teacher of TCAP-tested subject	-.041	.070	.840	-.21	.13
	Other Licensed Professional	.025	.082	.954	-.18	.23
Other Licensed Professional	Teacher of TCAP-tested subject	-.066	.074	.670	-.25	.12
	Teacher of non TCAP-tested subject	-.025	.082	.954	-.23	.18

Appendix Q

Pearson Chi-Square Test Results for Administrators by School Level Subgroup

Table Q1

Current Teacher Evaluations, Administrators by School Level

Label	Question:	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	–	–	–	–	–	–	–
C-TE_2	Current TE has provided me with meaningful feedback.	–	–	–	–	–	–	–
C-TE_3	Current TE has made me a better teacher overall.	–	–	–	–	–	–	–
C-TE_1a	Current TE encourages professional growth for teachers.	95%	79%	100%	67%	7.941	6	0.242
C-TE_2a	Current TE is linked to student learning.	81%	71%	86%	67%	4.021	6	0.674
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	91%	79%	100%	67%	21.196	9	0.012
C-TE_4a	Current TE is a high stress process for teachers.	38%	57%	69%	33%	10.478	9	0.313
C-TE_5a	Current TE is a high stress process for administrators.	29%	29%	43%	33%	7.016	9	0.635
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	71%	64%	79%	67%	27.140	9	0.001
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	67%	71%	86%	67%	22.296	9	0.008
	Counts	21	14	14	3	–	–	–

Note. Counts are an average for high school administrators. All questions had at least nine cells with an expected count less than five, which may have impacted significance.

Table Q2

Ideal Teacher Evaluations, Administrators by School Level

Label	Question: Ideal teacher evaluations:	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	86%	71%	79%	100%	11.935	9	.217
IdealTE_2	help teachers become better teachers.	95%	93%	100%	100%	2.770	6	.837
IdealTE_3	have the ability to improve student learning.	95%	100%	93%	100%	4.893	6	.558
IdealTE_4	are linked to student assessment data of any form.	81%	93%	79%	100%	10.730	6	.097
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	76%	57%	72%	100%	5.667	9	.773
IdealTE_6	are linked to student assessment data created by the teacher.	76%	86%	69%	100%	7.845	9	.550
	Counts	21	14	14	3	–	–	–

Note. Counts are an average for high school administrators. All questions had at least six cells with an expected count less than five, which may have impacted significance.

Table Q3

Education Reform, Administrators by School Level

Label	Question	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	86%	50%	86%	67%	12.419	9	.191
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	62%	36%	64%	67%	7.674	9	.567
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	67%	50%	71%	67%	3.127	9	.959
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	10%	0%	14%	0%	7.546	9	.580
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	10%	0%	14%	0%	3.130	9	.959
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	71%	57%	57%	100%	4.509	9	.875
	Counts	21	14	14	3	–	–	–

Note. All questions had at least 10 cells with an expected count less than five, which may have impacted significance.

Table Q4

SB191 Teacher Evaluations, Administrators by School Level

Label	Question	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	95%	86%	86%	67%	20.831	12	.053
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	57%	29%	29%	33%	18.874	12	.092
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	48%	21%	21%	33%	8.068	12	.780
	Counts	21	14	14	3	–	–	–

Note. All questions had at least 14 cells with an expected count less than five, which may have impacted significance.

Table Q5

SB 191 Teacher Evaluations, Administrators by School Level

Label	Question	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	95%	71%	79%	67%	6.199	6	.401
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	19%	7%	21%	33%	21.531	12	.043
SB191_3	I supported SB 191 during its development.	62%	43%	43%	33%	14.208	12	.288
SB191_4	I was pleased with the final version of SB 191 as it was passed.	71%	29%	57%	67%	21.139	9	.012
SB191_5	The media portrayed teachers fairly during the bill's passage.	29%	14%	23%	0%	7.056	12	.854
SB191_6	The media portrayed administrators fairly during the bill's passage.	43%	14%	23%	0%	9.980	9	.352
SB191_7	The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.	48%	46%	39%	0%	7.390	9	.597
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	95%	71%	79%	67%	20.850	9	.013
	Counts	21	14	14	3	—	—	—

Note. Counts are an average for high school administrators. All questions had at least nine cells with an expected count less than five, which may have impacted significance.

Table Q6

Change, Administrators by School Level

Label	Question	Elementary	Middle	High	K-8 + Other	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	100%	100%	100%	100%	7.096	3	.069
Change_2	I respond well when policies change at work.	100%	93%	100%	100%	6.162	6	.405
Change_3	I respond well to change at work when I am involved with creating the change.	100%	100%	100%	100%	4.836	3	.184
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	95%	100%	93%	100%	11.036	6	.087
Change_5	There is a right way and a wrong way to introduce new policies in a school.	91%	100%	93%	100%	12.333	6	.055
	Counts	21	14	14	3	–	–	–

Note. Counts are an average for high school administrators. All questions had at least three cells with an expected count less than five, which may have impacted significance.

Appendix R

Pearson Chi-Square Test Results for Administrators by Gender Subgroup

Table R1

Current Teacher Evaluations, Administrators by Gender

Label	Question:	Male	Female	Value	df	Asymp. Sig (2-Sided)
C-TE_1	Current TE improved your teaching.	–	–	–	–	–
C-TE_2	Current TE has provided me with meaningful feedback.	–	–	–	–	–
C-TE_3	Current TE has made me a better teacher overall.	–	–	–	–	–
C-TE_1a	Current TE encourages professional growth for teachers.	90%	91%	4.511	2	.105
C-TE_2a	Current TE is linked to student learning.	74%	82%	1.572	2	.456
C-TE_3a	Current TE is able to assess teachers' overall teaching abilities.	90%	88%	3.828	3	.281
C-TE_4a	Current TE is a high stress process for teachers.	61%	44%	3.377	3	.337
C-TE_5a	Current TE is a high stress process for administrators.	42%	26%	3.878	3	.275
C-TE_6a	Current TE allows for thorough reviews of teachers' overall teaching abilities.	63%	76%	2.163	3	.539
C-TE_7a	Current TE allows for thoughtful reviews of teachers' overall teaching abilities.	68%	77%	1.266	3	.737
	Counts	19	34	–	–	–

Note. Counts are an average. All questions had at least two cells with an expected count less than five, which may have impacted significance.

Table R2

Ideal Teacher Evaluations, Administrators by Gender

Label	Question: Ideal teacher evaluations:	Male	Female	Value	df	Asymp. Sig (2-Sided)
IdealTE_1	distinguish strong teachers from weak teachers.	74%	82%	4.557	3	.207
IdealTE_2	help teachers become better teachers.	100%	94%	2.557	2	.278
IdealTE_3	have the ability to improve student learning.	95%	97%	0.34	2	.843
IdealTE_4	are linked to student assessment data of any form.	95%	79%	3.799	2	.150
IdealTE_5	are linked to student assessment data by the state department of education (i.e., CSAPs).	79%	68%	1.843	3	.606
IdealTE_6	are linked to student assessment data created by the teacher.	90%	73%	5.493	3	.139
	Counts	19	34	–	–	–

Note. Counts are an average for female administrators. All questions had at least one cell with an expected count less than five, which may have impacted significance.

Table R3

Education Reform, Administrators by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
EdRef_1	EdRef: Education reforms improve student learning.	68%	79%	2.270	3	.518
EdRef_2	EdRef: Education reforms view teachers as knowledgeable professionals.	53%	59%	2.374	3	.498
EdRef_3	EdRef: Education reforms view building administrators as knowledgeable professionals.	63%	65%	1.332	3	.722
EdRef_4	EdRef: Policymakers, teachers have the same perceptions about what is needed to improve public education.	11%	6%	1.835	3	.607
EdRef_5	EdRef: Policymakers, building admins have the same perceptions about what is needed to improve public education.	11%	6%	4.742	3	.192
EdRef_6	EdRef: Teachers, building admins have the same perceptions about what is needed to improve public education.	58%	71%	2.846	3	.416
	Counts	19	34	–	–	–

Note. All questions had at least four cells with an expected count less than five, which may have impacted significance.

Table R4

Senate Bill 191 Teacher Evaluations, Administrators by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
SB191TE_1	SB 191's new TE: I will play an active role in the implementation of my school's new teacher evaluation system.	100%	82%	5.469	4	.242
SB191TE_2	SB 191's new TE: Teachers at my school believe it will be beneficial to the teachers' professional practice at my school.	42%	38%	1.788	4	.775
SB191TE_3	SB 191's new TE: Teachers at my school believe it will be beneficial to students at my school.	26%	35%	1.527	4	.822
	Counts	19	34	–	–	–

Note. All questions had at least five cells with an expected count less than five, which may have impacted significance.

Table R5

SB191 General, Administrators by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
SB191_1	I have a strong understanding of the legislation SB 191.	74%	88%	2.284	2	.319
SB191_2	I was highly involved in the development of SB 191 during its implementation phase.	26%	12%	4.073	4	.396
SB191_3	I supported SB 191 during its development.	63%	44%	6.618	4	.157
SB191_4	I was pleased with the final version of SB 191 as it was passed.	58%	56%	1.276	3	.735
SB191_5	The media portrayed teachers fairly during the bill's passage.	22%	21%	1.454	4	.835
SB191_6	The media portrayed administrators fairly during the bill's passage.	22%	29%	0.669	3	.880
SB191_7	The media portrayed policymakers (i.e.,legislators, lobbyists) fairly during the bill's passage.	33%	46%	0.824	3	.844
SB191_8	The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way.	84%	82%	1.752	3	.625
	Counts	19	34	–	–	–

Note. Counts are an average. All questions had at least two cells with an expected count less than five, which may have impacted significance.

Table R6

Change, Administrators by Gender

Label	Question	Male	Female	Value	df	Asymp. Sig (2-Sided)
Change_1	I view change as an opportunity for growth.	100%	100%	2.373	1	.123
Change_2	I respond well when policies change at work.	100%	97%	3.749	2	.153
Change_3	I respond well to change at work when I am involved with creating the change.	100%	100%	1.514	1	.218
Change_4	I respond well to change at work when informed of what new policies will be and I must learn the new policies.	100%	94%	6.799	2	.033
Change_5	There is a right way and a wrong way to introduce new policies in a school.	95%	94%	0.307	2	.857
	Counts	19	34	–	–	–

Note. Counts are an average for female administrators. Change_2, Change_4, and Change_5 had at least two cells with an expected count less than five, which may have impacted significance. There was a continuity correction for Change_1 and Change_3, for they were computed only for a 2x2 table.

Appendix S

One-Way ANOVA with post-hoc Scheffé Test Results for Administrators by School Level Subgroup

Table S1

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 47) = .991, p = .405

Current TE is able to assess teachers' overall teaching abilities. (CTE_3a)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.119	.207	.953	-.48	.72
	High School	-.183	.211	.861	-.80	.43
	K-8 + Other	.381	.370	.786	-.69	1.45
Middle School	Elementary School	-.119	.207	.953	-.72	.48
	High School	-.302	.231	.636	-.97	.37
	K-8 + Other	.262	.381	.924	-.84	1.37
High School	Elementary School	.183	.211	.861	-.43	.80
	Middle School	.302	.231	.636	-.37	.97
	K-8 + Other	.564	.384	.545	-.55	1.68
K-8 + Other	Elementary School	-.381	.370	.786	-1.45	.69
	Middle School	-.262	.381	.924	-1.37	.84
	High School	-.564	.384	.545	-1.68	.55

Table S2

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 48) = .937, p = .430

Current TE allows for thorough reviews of teachers' overall teaching abilities. (CTE_6a)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.071	.212	.990	-.54	.69
	High School	-.286	.212	.616	-.90	.33
	K-8 + Other	.048	.380	.999	-1.05	1.15
Middle School	Elementary School	-.071	.212	.990	-.69	.54
	High School	-.357	.233	.507	-1.03	.32
	K-8 + Other	-.024	.391	1.000	-1.16	1.11
High School	Elementary School	.286	.212	.616	-.33	.90
	Middle School	.357	.233	.507	-.32	1.03
	K-8 + Other	.333	.391	.867	-.80	1.47
K-8 + Other	Elementary School	-.048	.380	.999	-1.15	1.05
	Middle School	.024	.391	1.000	-1.11	1.16
	High School	-.333	.391	.867	-1.47	.80

Table S3

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 48) = .952, *p* = ..423

Current TE allows for thoughtful reviews of teachers' overall teaching abilities. (CTE_7a)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	-.071	.224	.992	-.72	.58
	High School	-.357	.224	.475	-1.01	.29
	K-8 + Other	.048	.401	1.000	-1.11	1.21
Middle School	Elementary School	.071	.224	.992	-.58	.72
	High School	-.286	.245	.717	-1.00	.43
	K-8 + Other	.119	.413	.994	-1.08	1.32
High School	Elementary School	.357	.224	.475	-.29	1.01
	Middle School	.286	.245	.717	-.43	1.00
	K-8 + Other	.405	.413	.811	-.79	1.60
K-8 + Other	Elementary School	-.048	.401	1.000	-1.21	1.11
	Middle School	-.119	.413	.994	-1.32	1.08
	High School	-.405	.413	.811	-1.60	.79

Table S4

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 48) = .516, p = .863

I was highly involved in the development of SB 191 during its implementation phase. (SB191_2)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.071	.498	.999	-1.37	1.52
	High School	-.357	.498	.915	-1.80	1.09
	K-8 + Other	0.000	.892	1.000	-2.58	2.58
Middle School	Elementary School	-.071	.498	.999	-1.52	1.37
	High School	-.429	.546	.892	-2.01	1.15
	K-8 + Other	-.071	.919	1.000	-2.73	2.59
High School	Elementary School	.357	.498	.915	-1.09	1.80
	Middle School	.429	.546	.892	-1.15	2.01
	K-8 + Other	.357	.919	.985	-2.31	3.02
K-8 + Other	Elementary School	0.000	.892	1.000	-2.58	2.58
	Middle School	.071	.919	1.000	-2.59	2.73
	High School	-.357	.919	.985	-3.02	2.31

Table S5

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 48) = 1.686, p = .183

I was pleased with the final version of SB 191 as it was passed. (SB191_4)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.500	.227	.197	-.16	1.16
	High School	.143	.227	.940	-.51	.80
	K-8 + Other	.048	.406	1.000	-1.13	1.22
Middle School	Elementary School	-.500	.227	.197	-1.16	.16
	High School	-.357	.248	.563	-1.08	.36
	K-8 + Other	-.452	.418	.761	-1.66	.76
High School	Elementary School	-.143	.227	.940	-.80	.51
	Middle School	.357	.248	.563	-.36	1.08
	K-8 + Other	-.095	.418	.997	-1.31	1.12
K-8 + Other	Elementary School	-.048	.406	1.000	-1.22	1.13
	Middle School	.452	.418	.761	-.76	1.66
	High School	.095	.418	.997	-1.12	1.31

Table S6

One-Way ANOVA with Scheffé Post-Hoc Test

F (3, 48) = 2.931, p = .043

The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way. (SB191_8)

(I) Administrators by School Level	(J) Administrators by School Level Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Elementary School	Middle School	.738*	.249	.044	.02	1.46
	High School	.310	.249	.675	-.41	1.03
	K-8 + Other	.381	.446	.866	-.91	1.67
Middle School	Elementary School	-.738*	.249	.044	-1.46	-.02
	High School	-.429	.273	.489	-1.22	.36
	K-8 + Other	-.357	.460	.895	-1.69	.98
High School	Elementary School	-.310	.249	.675	-1.03	.41
	Middle School	.429	.273	.489	-.36	1.22
	K-8 + Other	.071	.460	.999	-1.26	1.40
K-8 + Other	Elementary School	-.381	.446	.866	-1.67	.91
	Middle School	.357	.460	.895	-.98	1.69
	High School	-.071	.460	.999	-1.40	1.26

*. The mean difference is significant at the 0.05 level.

Appendix T

One-Way ANOVA with Scheffé Post-Hoc Test Results for Teachers, Administrators, and Policymakers

Table T1

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 602) = 6.108, p = .002

Ideal TE distinguish strong teachers from weak teachers. (IdealTE_1)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.336*	.123	.025	-.64	-.03
	Policymaker	-.686	.299	.073	-1.42	.05
Administrator	Teacher	.336*	.123	.025	.03	.64
	Policymaker	-.351	.319	.547	-1.13	.43
Policymaker	Teacher	.686	.299	.073	-.05	1.42
	Administrator	.351	.319	.547	-.43	1.13

*. The mean difference is significant at the 0.05 level.

Table T2

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 606) = 12.889 , p = .000

Ideal TE are linked to student assessment data of any form. (IdealTE_4)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.533*	.109	.000	-.80	-.27
	Policymaker	-.403	.265	.315	-1.05	.25
Administrator	Teacher	.533*	.109	.000	.27	.80
	Policymaker	.130	.282	.899	-.56	.82
Policymaker	Teacher	.403	.265	.315	-.25	1.05
	Administrator	-.130	.282	.899	-.82	.56

*. The mean difference is significant at the 0.05 level.

Table T3

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 605) = 21.896, p = .000

Ideal TE are linked to student assessment data by the state department of education (i.e., CSAPs). (IdealTE_5)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.749*	.114	.000	-1.03	-.47
	Policymaker	-.221	.276	.727	-.90	.46
Administrator	Teacher	.749*	.114	.000	.47	1.03
	Policymaker	.529	.295	.201	-.19	1.25
Policymaker	Teacher	.221	.276	.727	-.46	.90
	Administrator	-.529	.295	.201	-1.25	.19

*. The mean difference is significant at the 0.05 level.

Table T4

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 590) = 4.654, p = .010

Education reforms improve student learning. (EdRef_1)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.224*	.088	.039	-.44	-.01
	Policymaker	-.367	.208	.211	-.88	.14
Administrator	Teacher	.224*	.088	.039	.01	.44
	Policymaker	-.143	.222	.812	-.69	.40
Policymaker	Teacher	.367	.208	.211	-.14	.88
	Administrator	.143	.222	.812	-.40	.69

*. The mean difference is significant at the 0.05 level.

Table T5

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 591) = 9.824, p = .000

Education reforms view teachers as knowledgeable professionals. (EdRef_2)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.411*	.103	.000	-.66	-.16
	Policymaker	-.506	.245	.120	-1.11	.10
Administrator	Teacher	.411*	.103	.000	.16	.66
	Policymaker	-.095	.262	.936	-.74	.55
Policymaker	Teacher	.506	.245	.120	-.10	1.11
	Administrator	.095	.262	.936	-.55	.74

*. The mean difference is significant at the 0.05 level.

Table T6

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 593) = 14.052, p = .000

Policymakers, teachers have the same perceptions about what is needed to improve public education.(EdRef_4)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.397*	.081	.000	-.60	-.20
	Policymaker	-.446	.195	.074	-.92	.03
Administrator	Teacher	.397*	.081	.000	.20	.60
	Policymaker	-.050	.208	.972	-.56	.46
Policymaker	Teacher	.446	.195	.074	-.03	.92
	Administrator	.050	.208	.972	-.46	.56

*. The mean difference is significant at the 0.05 level.

Table T7

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 593) = 5.269, p = .005

Policymakers, building admins have the same perceptions about what is needed to improve public education.(EdRef_5)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.246*	.086	.017	-.46	-.04
	Policymaker	-.335	.205	.264	-.84	.17
Administrator	Teacher	.246*	.086	.017	.04	.46
	Policymaker	-.089	.219	.920	-.63	.45
Policymaker	Teacher	.335	.205	.264	-.17	.84
	Administrator	.089	.219	.920	-.45	.63

*. The mean difference is significant at the 0.05 level.

Table T8

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 5982) = 13.503, p = .000

I have a strong understanding of the legislation SB 191. (SB191_1)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.582*	.123	.000	-.88	-.28
	Policymaker	-.674	.289	.067	-1.38	.04
Administrator	Teacher	.582*	.123	.000	.28	.88
	Policymaker	-.093	.310	.956	-.85	.67
Policymaker	Teacher	.674	.289	.067	-.04	1.38
	Administrator	.093	.310	.956	-.67	.85

*. The mean difference is significant at the 0.05 level.

Table T9

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 583) = 7.812, p = .000

I was highly involved in the development of SB 191 during its implementation phase. (SB191_2)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.462*	.171	.027	-.88	-.04
	Policymaker	-1.202*	.402	.012	-2.19	-.22
Administrator	Teacher	.462*	.171	.027	.04	.88
	Policymaker	-.741	.430	.228	-1.80	.32
Policymaker	Teacher	1.202*	.402	.012	.22	2.19
	Administrator	.741	.430	.228	-.32	1.80

*. The mean difference is significant at the 0.05 level.

Table T10

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 581) = 5.668, p = .004

I supported SB 191 during its development. (SB191_3)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.505*	.202	.044	-1.00	-.01
	Policymaker	-1.116	.475	.064	-2.28	.05
Administrator	Teacher	.505*	.202	.044	.01	1.00
	Policymaker	-.611	.509	.486	-1.86	.64
Policymaker	Teacher	1.116	.475	.064	-.05	2.28
	Administrator	.611	.509	.486	-.64	1.86

*. The mean difference is significant at the 0.05 level.

Table T11

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 579) = 2.020, p = .134

I was pleased with the final version of SB 191 as it was passed. (SB191_4)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.327	.192	.235	-.80	.14
	Policymaker	-.512	.451	.526	-1.62	.59
Administrator	Teacher	.327	.192	.235	-.14	.80
	Policymaker	-.185	.483	.929	-1.37	1.00
Policymaker	Teacher	.512	.451	.526	-.59	1.62
	Administrator	.185	.483	.929	-1.00	1.37

Table T12

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 576) = 1.417, *p* = ..243

The media portrayed teachers fairly during the bill's passage. (SB191_5)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.321	.191	.244	-.79	.15
	Policymaker	-.071	.445	.987	-1.16	1.02
Administrator	Teacher	.321	.191	.244	-.15	.79
	Policymaker	.249	.477	.872	-.92	1.42
Policymaker	Teacher	.071	.445	.987	-1.02	1.16
	Administrator	-.249	.477	.872	-1.42	.92

Table T13

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 574) = .447, p = .640

The media portrayed administrators fairly during the bill's passage. (SB191_6)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.163	.197	.711	-.65	.32
	Policymaker	-.226	.459	.886	-1.35	.90
Administrator	Teacher	.163	.197	.711	-.32	.65
	Policymaker	-.063	.493	.992	-1.27	1.15
Policymaker	Teacher	.226	.459	.886	-.90	1.35
	Administrator	.063	.493	.992	-1.15	1.27

Table T14

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 576) = 7.768, p = .000

The reforms coming out of SB 191 have a chance to improve Colorado's students' learning in a meaningful way. (SB191_8)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.604*	.161	.001	-1.00	-.21
	Policymaker	-.512	.378	.401	-1.44	.42
Administrator	Teacher	.604*	.161	.001	.21	1.00
	Policymaker	.093	.405	.974	-.90	1.09
Policymaker	Teacher	.512	.378	.401	-.42	1.44
	Administrator	-.093	.405	.974	-1.09	.90

*. The mean difference is significant at the 0.05 level.

Table T15

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 603) = 8.853, p = .000

I view change as an opportunity for growth (Change_1).

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.300*	.077	.001	-.49	-.11
	Policymaker	-.329	.188	.216	-.79	.13
Administrator	Teacher	.300*	.077	.001	.11	.49
	Policymaker	-.029	.200	.990	-.52	.46
Policymaker	Teacher	.329	.188	.216	-.13	.79
	Administrator	.029	.200	.990	-.46	.52

*. The mean difference is significant at the 0.05 level.

Table T16

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 601) = 14.127, p = .000

I respond well when policies change at work. (Change_2)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.416*	.079	.000	-.61	-.22
	Policymaker	-.140	.191	.766	-.61	.33
Administrator	Teacher	.416*	.079	.000	.22	.61
	Policymaker	.276	.204	.400	-.22	.78
Policymaker	Teacher	.140	.191	.766	-.33	.61
	Administrator	-.276	.204	.400	-.78	.22

*. The mean difference is significant at the 0.05 level.

Table T17

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 603) = 11.148, p = .000

I respond well to change at work when informed of what new policies will be and I must learn the new policies. (Change_4)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	-.427*	.093	.000	-.66	-.20
	Policymaker	.226	.227	.610	-.33	.78
Administrator	Teacher	.427*	.093	.000	.20	.66
	Policymaker	.653*	.242	.027	.06	1.25
Policymaker	Teacher	-.226	.227	.610	-.78	.33
	Administrator	-.653*	.242	.027	-1.25	-.06

*. The mean difference is significant at the 0.05 level.

Table T18

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 562) = 9.343, p = .000

Central focus of SB 191 (CF)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	.562*	.144	.001	.21	.91
	Policymaker	.710	.355	.136	-.16	1.58
Administrator	Teacher	-.562*	.144	.001	-.91	-.21
	Policymaker	.149	.378	.926	-.78	1.08
Policymaker	Teacher	-.710	.355	.136	-1.58	.16
	Administrator	-.149	.378	.926	-1.08	.78

*. The mean difference is significant at the 0.05 level.

Table T19

One-Way ANOVA with Scheffé Post-Hoc Test

F (2, 589) = 1.709, p = .182

Primary responsibility for successful implementation of SB 191? (PR)

(I) TAP	(J) TAP Comparison	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Teacher	Administrator	.327	.200	.265	-.16	.82
	Policymaker	-.383	.475	.723	-1.55	.78
Administrator	Teacher	-.327	.200	.265	-.82	.16
	Policymaker	-.709	.508	.378	-1.96	.54
Policymaker	Teacher	.383	.475	.723	-.78	1.55
	Administrator	.709	.508	.378	-.54	1.96